DURATIONAL CUES TO WORD BOUNDARIES IN CLEAR SPEECH

Anne CUTLER and Sally BUTTERFIELD

MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK

Abstract. One of a listener's major tasks in understanding continuous speech is segmenting the speech signal into separate words. When listening conditions are difficult, speakers can help listeners by deliberately speaking more clearly. In four experiments, we examined how word boundaries are produced in deliberately clear speech. Durational measurements were taken of the pre-boundary syllable, and of pausing (if any) at the boundary, in baseline utterances and in deliberately clear repetitions. We found that speakers do indeed attempt to mark word boundaries in clear (though not in normal) speech; moreover, they differentiate between word boundaries in a way which suggests they are sensitive to listener needs. Previous research has suggested that in English, listeners apply heuristic segmentation strategies which make word boundaries before strong syllables easiest to perceive. When conditions for the listener are difficult, however, speakers pay more attention to marking word boundaries before weak syllables, i.e. they mark just those boundaries which are otherwise particularly hard to perceive.

Zusammenfassung. Eine der Hauptaufgaben fur den Horer beim Verstehen fliessender Sprache ist die. Zerlegung des Sprachsignales in Worter. Wenn die Horbedingungen schwierig sind, kann der Sprecher dem Zuhorer durch bewuBt deutliches Sprechen helfen. In vier Experimenten haben wir untersucht, wie Wortgrenzen in bewuBt deutlich gesprochener Sprache erzeugt werden. Es wurden fur einfache und bewuBt deutliche Aussprache sowohl die Silbenlangen vor Wortgrenzen bestimmt, als auch die Pausen zwischen den Wortern (wenn dort welche waren). Wir fanden, daB Sprecher bei klarer Sprechweise (nicht bei normaler) tatsachlich versuchen, Wortgrenzen zu markieren; weiter wird zwischen einzelnen Wortgrenzen auf einer Art unterschieden, die eine Empfindlichkeit des Sprechers gegenuber die Bedurfnisse des Horers aufweist. Vorhergehende Forschung hat uns schliessen lassen, daB in Englisch die Zuhorer heuristische Unterteilungsstrategien anwenden. Diese fiihren dazu, daB Wortgrenzen vor starken Silben am einfachsten wahrzunehmen sind. Wenn die Bedingungen fur den Horer schwierig sind, achten Sprecher besonders darauf, Wortgrenzen vor schwachen Silben zu betonen; das heiBt, sie markieren gerade die Grenzen, die sonst schwer zu erkennen waren.

Resume. L'une des taches majeures de l'auditeur lors du processus de comprehension de la parole est la segmentation du signal en mots. Quand les conditions d'ecoute sont mauvaises, les locuteurs peuvent aider ceux qui les ecoutent en articulant clairement. Nous avons examine, au cours de quatre experiences, la production des frontieres de mots dans la parole d£liberement claire. Des mesures de la duree des syllabes qui precedent les frontieres de mots, ainsi que celle de pauses eventuelles ont ete faites dans une serie de phrases prononcees normalement et dans ces memes phrases prononcees de maniere particulierement claire. Nous avons trouve que quand les locuteurs articulent tres soigneusement, ils essaient effectivement de souligner les frontieres de mots; de plus, la maniere dont ils s'y prennent suggere qu'ils sont sensibles aux besoins de ceux qui les ecoutent. En effet, des etudes precedentes ont suggere que les Anglais emploient une strategic de segmentation telle qu'ils reperent plus aisement les frontieres de mots precedant les syllabes accentuees. Or, dans des conditions d'ecoute difficiles, les locuteurs s'appliquent a souligner les frontieres de mots precedant les syllabes faibles, c'est-a-dire celles-la memes qui seraient particulierement difficiles a percevoir.

Keywords. Speech production, intelligibility, clear speech, word boundaries, segmentation, duration, lengthening, pausing.

1. Introduction

To understand continuous speech, a recogniser has to locate and identify parts of the speech signal which correspond to individual words. Unfortunately, segmenting continuous speech into words is not easy, since word boundaries are seldom explicitly marked. Human listeners respond to this problem by adopting various strategies to maximise the efficiency of word boundary loca-

tion. For instance, an efficient strategy for segmenting continuous English speech would be to assume that strong syllables, i.e. syllables in which the vowel quality of the nucleus is full rather than reduced, are highly likely to be word-initial, because word counts of spontaneously produced British English speech have shown that about 90% of lexical words (content words) begin with strong syllables (Cutler and Carter, 1987). Indeed, experimental studies have shown that listeners do segment English speech at the onset of strong syllables (Cutler and Norris, 1988). Moreover, when listeners misperceive word boundaries, their most likely mistake is the erroneous insertion of a boundary before a strong syllable (Butterfield and Cutler, 1988).

However, human speakers can, if necessary, make word boundaries clear. For instance, a speaker could pause before every word. As we know, conversational speech is *never* like that. But speakers do speak in a range of styles, using careful articulation with foreigners, for example, but casual mumbles with close friends and family. And several recent studies have demonstrated that speakers who notice that a listener is having difficulty do indeed adjust their speech towards clearer articulation when repeating. Adjustments include speaking more slowly, louder, and with raised pitch (Clark et al., 1988); making syntactic structure explicit (Valian and Wales, 1976); and implementing segmental changes such as separating the VOT distributions for voiced and voiceless stop consonants and fully releasing word-final stops (Chen et al., 1983; Picheny et al., 1986).

Few such studies have examined precisely how word boundaries are produced when speakers are deliberately trying to speak clearly. Picheny et al., (1986) found that clear speech contains pauses at word boundaries, although most such pauses were much shorter than the 250 ms which is commonly used as the threshold for defining a pause in other studies (Grosjean, 1980). They did not, in their report, distinguish between *types* of word boundaries. Studies of normal speech production, however, suggest that not all word boundaries will necessarily be treated equally. Cooper and Paccia-Cooper (1980) studied the application across word boundaries of phonological rules such as palatalisation, and found that speakers are reluc-

tant to apply such rules when they will distort the initial boundaries of low frequency or contrastively stressed words; however, they are happy to apply them across the initial boundaries of high frequency, unstressed words.

Given that the evidence from recognition studies suggests that the distinction between strong and weak word-initial syllables is important for speech segmentation, it is reasonable to ask whether this distinction is also relevant in describing clear speech phenomena. When speakers know that listening conditions are difficult, they may pay particular attention to helping listeners with the segmentation problem, by trying hard to make word boundaries clear; moreover, they may distinguish between types of word boundaries by making some even clearer than others. In the present study we investigated whether such a distinction is made in clear speech between boundaries preceding strong versus weak syllables.

The Cooper and Paccia-Cooper findings for natural speech point to several factors which might affect the strong/weak comparison. Word frequency, which Cooper and Paccia-Cooper found to predict initial boundary distortions, is related to prosodic structure of initial syllables. The relationship, however, is complex. Among open class words (nouns, verbs and adjectives), words with weak initial syllables tend to occur less frequently than words with strong initial syllables. Closed class (or function) words, however, tend almost with exception to be prosodically weak but to have a very high frequency of occurrence. Contrastive stress, which Cooper and Paccia-Cooper found would block initial boundary distortions, also affects strong and weak syllables differently. Furthermore, Cooper and Paccia-Cooper suggested that the frequency and stress effects might be separate manifestations of a single underlying effect whereby the initial boundaries of *unpredictable* words resist distortion: low frequency words tend to be less predictable in context than high frequency words, and contrastive stress tends to be assigned to unpredictable words. We therefore paid attention to these factors as well as to the primary comparison between strong and weak syllables.

2. Experiment 1

2.1. Method

2.1.1. Materials

Twelve sentences of relatively unpredictable content were constructed. Each sentence contained a critical word boundary; in six sentences the word after this boundary began with a strong syllable, in six it began with a weak syllable. The sentences were paired so that phonetic material immediately either side of the boundary was comparable in a strong-syllable and a weak-syllable case. Examples are "Take it in turns to eat breakfast", where the critical boundary precedes "turns" (a strong syllable), versus "He called in to view it himself, where the critical boundary precedes "to" (here, a weak syllable).

The form of the sentences was in part determined by the availability of possible mishearings in which the critical boundary was absent. For instance, "in turns" could be misheard as "interns", while "in to" could be misheard as "into". For each sentence we constructed two such purported mishearings, to be presented to the subjects as feedback. These were quite realistic as mishearings - the rhythm of the sentence was fairly well preserved, as were most of the vowels in the stressed syllables. In each case, however, the feedback sentences contained *no* boundary at the critical location. For the above examples, the feedback sentences were "Baker interns all the terrorists" and "Take it internally at breakfast", versus "The cold interviewer was selfish" and "He crawled into view by himself. The full set of target and feedback sentences is listed in the Appendix.

The members of each pair were compared on word frequency (using the Francis and Kucera, 1982, word-class-specific norms) and predictability measures. Word class was not matched within sentence pairs, and as a result frequency was also mismatched: five of the six weak syllables were high frequency closed class words. Thus the frequency of occurrence for the weak syllables was in general much higher than that of the strong syllables. Predictability was assessed by means of a cloze test: 40 native English speakers not otherwise involved in the experiments were presented

with the target sentences from which the words after the critical boundary had been removed (e.g. "Take it in_____to eat breakfast"; "He called in_____view it himself), and were asked to supply a single word completion. The predictability of the weak syllables (mean 32%) was somewhat higher than that of the strong syllables (mean 19%), but this was chiefly due to 100% predictability of "to" in the above example.

2.1.2. Subjects and procedure

Five members of the Applied Psychology Unit subject panel took part (for payment) in the experiment. They were told that their speech was being fed through a distorting filter to a listener in the next room who would type what he thought he heard into a computer which in turn would display this response on the subjects' VDU screen. In fact the only listener was the experimenter, and all subjects received the same "feedback". The subjects were given the sentences on cards, and were asked to read each sentence as naturally as possible when first producing it. If the listener's response was incorrect, then the sentence should be repeated; if the second response was again incorrect, the sentence should be repeated once more. Because for each experimental sentence the "listener's" response was indeed twice incorrect, this instruction ensured that these sentences were produced three times each. The subjects were asked to speak clearly when repeating (but they were told not to shout as this would make the distortion worse).

Besides the 12 experimental sentences, subjects produced three practice and ten filler sentences, some of which the "listener" apparently heard correctly on first or second hearing. All the subjects' productions were recorded.

Each subject's three utterances of each of the 12 sentences were digitised - 180 utterances in all. This rich body of data lends itself to a variety of analyses. With respect solely to the critical word boundaries, we could examine whether there is pausing at boundaries; whether pre-boundary syllables are lengthened; whether intonational movement marks boundaries; whether word-initial or word-final segments are articulated differently in first versus later utterances; etc. Further analyses will be described in sub-

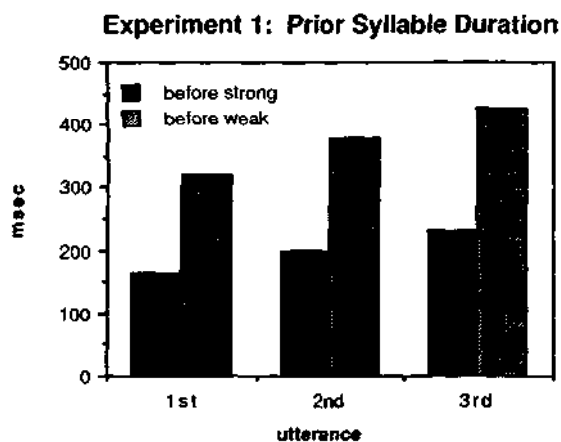**Experiment 1: Prior Syllable Duration**

Fig. 1. Duration (ms) of the syllable preceding the boundary in the baseline, first repetition and second repetition, separately for boundaries before strong and before weak syllables in Experiment 1.

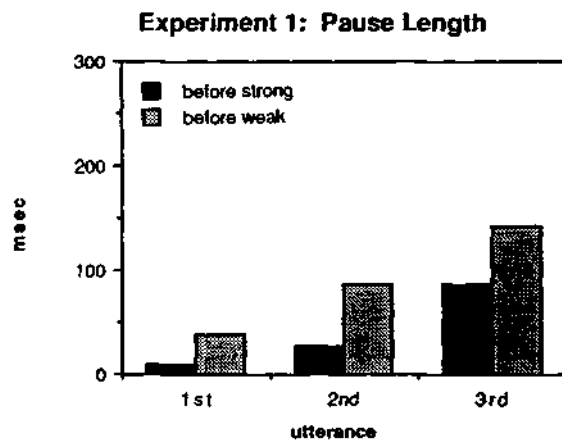**Experiment 1: Pause Length**

Fig. 2. Duration (ms) of pauses at the boundary in the baseline, first repetition and second repetition, separately for boundaries before strong and before weak syllables in Experiment 1.

sequent reports; the present report describes just two durational measures (of pauses and of pre-boundary syllables). The measurements were carried out on the digitised waveforms using the CAMSED speech editing package.

2.2. Results and discussion

Figures 1 and 2 show the mean durations (across subjects and sentences) of the pauses and pre-boundary syllables in first, second and third productions, separately for utterances where the boundary preceded a strong versus a weak syllable.

It can be seen that pre-boundary syllables preceding weak word-initial syllables were longer than those preceding strong syllables, even in the baseline utterances. This is an artefact of our materials; English has a tendency to alternate weak and strong syllables, and in some of our sentences weak syllables were preceded by strong syllables, and vice versa. Thus the proper measure on which to base statistical analysis is the amount of lengthening from first to second utterance, and from second to third. Analyses of variance were conducted with both subjects (F_1) and sentences (F_2) as random factors; although greater lengthening occurred before weak than

before strong syllables, the differences failed to reach conventional levels of statistical significance $(F_1 [1,4] = 4.05, p > 0.1; F_2 [1,10] = 4.24, p < 0.07)$.

Pauses, however, were significantly longer before weak syllables $(F_1 [1,4] = 2.63, p > 0.1$, but $F_2[1,10] = 5.91, p < 0.04)$.

Curiously, however, the pause results reveal no interaction between the effect of strong versus weak syllables and the effect of repetitions; as is clear from the figures, pauses were longer at pre-weak than at pre-strong boundaries even in the baseline utterances. This would seem to suggest that these differences may be characteristic of normal speech production, and not specific to deliberately clear speech; but such effects have not previously been reported in studies of normal speech production, so that this suggestion seems unlikely. An alternative possibility is that subjects were speaking clearly even in the baseline utterances. In Experiment 2, therefore, we used as a baseline utterances which were collected before subjects were aware of the need to speak clearly.

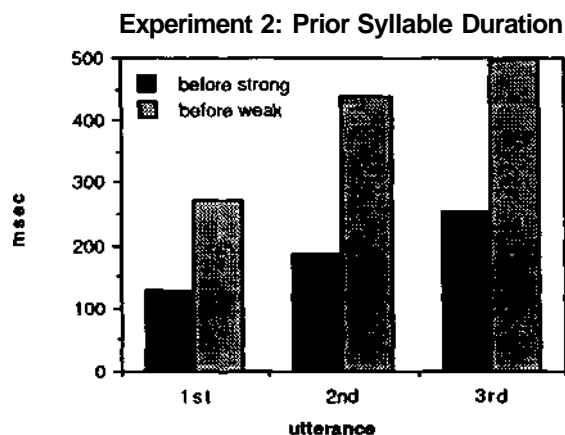## Experiment 2: Prior Syllable Duration

Fig. 3. Duration (ms) of the syllable preceding the boundary in the baseline, first repetition and second repetition, separately for boundaries before strong and before weak syllables in Experiment 2.
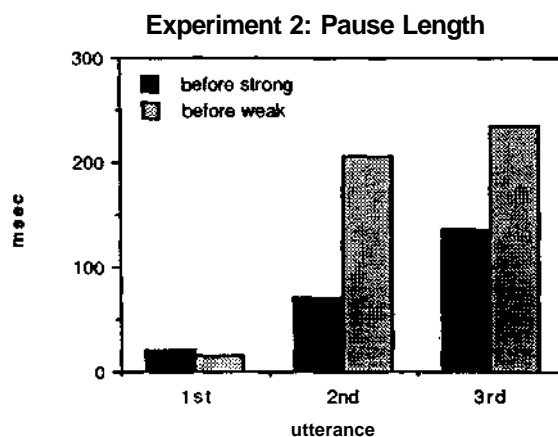
Experiment 2: Pause Length

Fig. 4. Duration (ms) of pauses at the boundary in the baseline, first repetition and second repetition, separately for boundaries before strong and before weak syllables in Experiment 2.

3. Experiment 2

3.1. Method

Five further subjects from the same population produced the same sentences under the same experimental conditions, with one exception: before being told about the listener and the supposed distortion, the subjects read the experimental (and filler) sentences aloud onto tape. These initial productions then served as the baseline to be measured and compared with the two repetitions after feedback.

3.2. Results and discussion

Figures 3 and 4 display the mean durations (across subjects and sentences) of the pauses and pre-boundary syllables in the baseline, second and third productions, separately for utterances where the boundary preceded a strong versus a weak syllable.

Again it can be seen that syllables preceding weak word-initial syllables were lengthened to a relatively greater degree than syllables preceding strong word-initial syllables, and the difference in increase this time was statistically significant $(F_1$ [1,4] = 12.54, $p < 0.03$; F_2 [1,10] = 7.85, $p <$

0.02). However, the effect of strong versus weak syllables interacted significantly with the effect of repetitions (F_1 [1,4] = 8.94, $p < 0.05$; F_2 [1,10] = 5.66, $p < 0.04$), and subsequent t-tests showed that the increase from baseline to second utterance was significantly greater before a pre-weak than before a pre-strong boundary (t [4] = 5.95, $p < 0.005$), but there was no significant difference between the two conditions in the increase from second to third utterance.

Pausing was again longer before a weak syllable (F_1 [1,4] = 10.93, $p < 0.03$; F_2 [1,10] = 2.56, $p > 0.1$), but the effects differed across repetitions: t-tests showed no difference in the baseline condition, but longer pausing before weak than before strong syllables which although it was not significant in the first repetition (t [4] = 2.04, $p = 0.11$), was highly significant in the second repetition (t [4] = 5.09, $p < 0.007$).

Both Experiments 1 and 2 have suggested that speakers who are deliberately trying to speak clearly do indeed produce cues to the presence of a word boundary; moreover, these cues tend to be more marked before a weak than before a strong initial syllable. However, it is possible that some differences between our sentence pairs might have contributed to the effects we found. For instance, as we pointed out above, the

pre-boundary syllables were imperfectly matched, so that some differences in pre-boundary syllable duration may reflect merely differences in pre-boundary syllable identity. In addition, in some cases there was a difference between the two members of a pair in the syntactic strength of the crucial boundary (although the differences involved minor phrase boundaries, which are not usually marked in normal speech production; cf. Cooper et al., 1978). Finally, as we pointed out above, word class of the word following the boundary was mismatched, as a result of which the weak syllables had a higher frequency than the strong syllables, as well as, to some extent, greater predictability. These latter two differences, it should be pointed out, are in the reverse direction to the claims made by Cooper and Paccia-Cooper (1980) for natural speech: the weak syllables, which received the greater marking, had _higher_ frequency and were _more_ predictable.

Two further experiments were conducted in an attempt to distinguish between various possible explanations for our findings. In Experiment 3, we manipulated prosodic structure while keeping word class constant, and in Experiment 4 we manipulated word class - the closed/open distinction - while keeping prosodic structure as far as possible constant. Because closed class words are always of higher average frequency than open class words, the frequency difference in Experiment 4 was in the same direction as in Experiments 1 and 2, and therefore in Experiment 3 we manipulated frequency in the opposite direction - all initially-strong words were of higher frequency than their initially-weak pairs. Predictability was factored out by matching this factor across the critical comparison in both experiments.

The combined results of Experiments 3 and 4 will allow us to decide between word class, word predictability, word frequency and prosodic structure as explanations of the findings of Experiments 1 and 2. (a) If the greater marking for weak syllables in Experiments 1 and 2 was because they were closed class words, then we will find greater marking for closed class words again in Experiment 4 but no difference between the conditions in Experiment 3. (b) If predictability was the operant factor in Experiments 1 and 2, then we will find no difference in either Experiment 3 or 4

since predictability is matched, (c) If the greater marking for weak syllables in Experiments 1 and 2 was because they had higher word frequency, then we will find greater marking for closed class words again in Experiment 4, but greater marking for strong-initial than for weak-initial words in Experiment 3. (d) Finally, if the results of Experiments 1 and 2 indeed reflected differences in prosodic structure, then we will find greater marking for weak than for strong initial syllables in Experiment 3, but (to the extent that we can succeed in controlling prosodic structure) no difference between the conditions in Experiment 4.

4. Experiment 3

4.1. Method

4.1.1. Materials

A further set of 12 sentences was constructed, again in six matched pairs containing strong and weak syllables after the critical boundary. Word class of the word after the boundary was matched in each pair, as was syntactic strength of the boundary and identity of the pre-boundary syllable. An example pair is "Play this card a good deal more'/'Fire this cadet's automatic"; the crucial boundary is "this c-". All the words with strong initial syllables were higher in frequency of occurrence than their weak-initial pairs. Predictability of the critical words was equated; cloze test results from the same 40 subjects yielded virtually zero predictability for all items in this set. Purported mishearings were again constructed for use as feedback. The complete set of target and feedback sentences is listed in the Appendix.

4.7.2. Subjects and procedure

Ten subjects from the same population took part; the procedure was as in Experiment 2. This experiment and the next were administered together, so that the total number of items, including the three practice and ten filler sentences, was 37 rather than the 25 of the preceding studies.

4.2. Results and discussion

Figures 5 and 6 show the mean durations (across subjects and sentences) of the pauses and
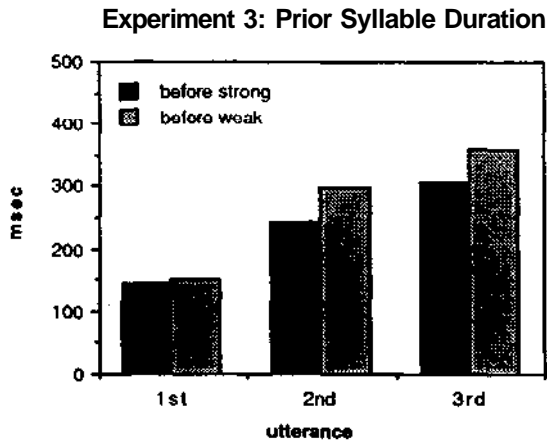
Experiment 3: Prior Syllable Duration

Fig. 5. Duration (ms) of the syllable preceding the boundary in the baseline, first repetition and second repetition, separately for boundaries before strong and before weak syllables in Experiment 3.
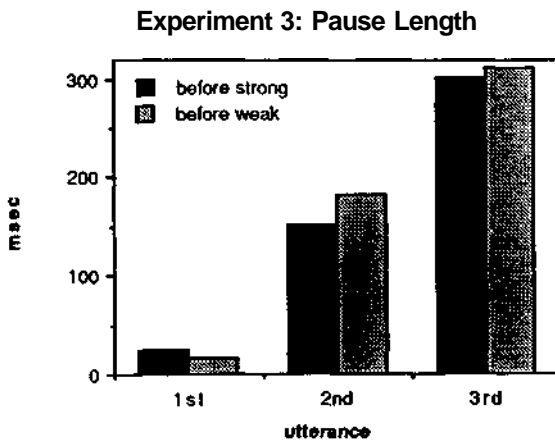
Experiment 3: Pause Length

Fig. 6. Duration (ms) of pauses at the boundary in the baseline, first repetition and second repetition, separately for boundaries before strong and before weak syllables in Experiment 3.

pre-boundary syllables in the baseline, second and third productions, separately for utterances where the boundary preceded a strong versus a weak syllable.

Statistical evaluation of the results for the pre-boundary syllables showed a similar picture to that found in Experiment 2. The duration of the pre-boundary syllable (recall that in this experi-

ment this syllable is the same in each member of a pair) hardly differed in the baseline, but was lengthened by a much greater amount before weak syllables than before strong in the repetitions; t-tests again showed that the increase from baseline to second utterance was significantly greater before weak than before strong syllables ($t [9] = 7.75, p < 0.001$), but there was no significant difference between the conditions in the amount of increase from second to third utterance.

Likewise, the pauses showed no difference in the baseline utterance, but in both repetitions pauses preceding weak syllables were longer than pauses preceding strong syllables (although in this case the effect failed to reach statistical significance).

The results of this experiment therefore again showed greater marking for word boundaries preceding weak than preceding strong syllables, as found in Experiments 1 and 2, although the effects were somewhat less robust than in the earlier studies.

5. Experiment 4

5.1. Method

5.1.1. Materials

A further set of 12 sentences was constructed, again in six matched pairs. In this case the critical variable was word class of the word after the boundary; homophones were chosen which could be either open class or closed class words. An example pair is *hour/our,* as in "Lots of hour-long sessions are needed" versus "Both of our children like peanuts"; the crucial boundary is "of (h)our". Although it would have been desirable to vary word class fully independently of the strong/weak syllable distinction, this is impossible because nearly all closed class words, but no open class words, can be reduced in sentence contexts; most closed class words which cannot be reduced - *these, those,* etc. - are not homophonous with open class words. Where we could, we chose homophones which could not be reduced, and for the remaining items we attempted to construct contexts in which reduction of the closed class

words was certainly not mandatory, and was relatively unlikely.

Each closed class word was higher in frequency than its open class pair. Predictability was fairly well matched, with predictability ratings for both the closed class (mean 25%) and the open class set (15%) being low to moderate. Purported mishearings were again constructed for use as feedback. The complete set of target and feedback sentences is listed in the Appendix.

5.1.2. Subjects and procedure

This experiment was administered together with Experiment 3.

5.2. Results and discussion

Figures 7 and 8 show the mean durations (across subjects and sentences) of the pauses and pre-boundary syllables in the baseline, second and third productions, separately for utterances where the boundary preceded an open versus a closed class word.

The duration of the pre-boundary syllable increased from baseline to first and again to second repetition, but the increases did not differ significantly for boundaries preceding closed versus open class words. Therefore the closed-open class comparison completely failed to replicate the significant difference in preceding syllable duration observed in all three preceding experiments with weak/strong syllable comparisons.

The duration of pauses did however show a difference between closed and open class boundaries; in the subjects analysis, the greater length of pauses before closed than before open class words was significant $(F_1 [1,9] = 6.12, p < 0.04)$. This effect interacted with the effect of repetitions $(F_1 [2,18] = 6.57, p < 0.01)$, and subsequent t tests across subjects showed that the difference was insignificant in the baseline and first repetition, but significant in the second repetition $(t [9] = 3.12, p < 0.02)$.

None of these effects reached significance in the items analysis, however, which suggests that there was considerable variability between the item pairs. Indeed, inspection of the item means revealed that all the difference was due to exceptionally long pauses, in the second repetition

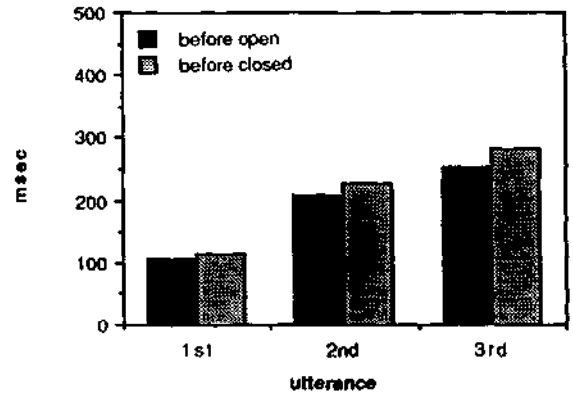**Experiment 4: Prior Syllable Duration**

Fig. 7. Duration of the syllable preceding the boundary in the baseline, first repetition and second repetition, separately for boundaries before closed class and before open class words in Experiment 4.
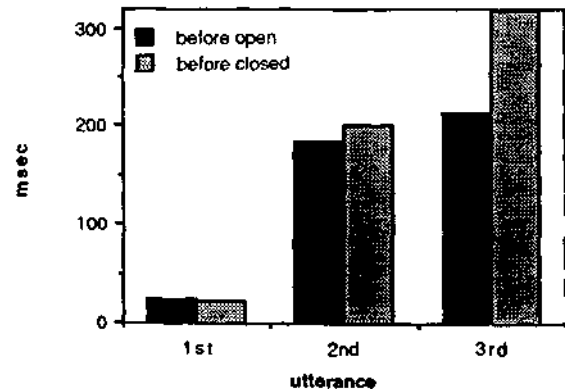
Experiment 4: Pause Length

Fig. 8. Duration of pauses at the boundary in the baseline, first repetition and second repetition, separately for boundaries before closed class and before open class words in Experiment 4.

only, before just three closed class words: *by, just* and *for*. Interestingly, in the sentence contexts which we created each of these words at least offers the possibility of vowel reduction, whereas of the three pairs which did not show greater pause duration before closed class words, two such words *(not, I)* virtually exclude the possibility of vowel reduction. Thus the apparent difference between closed and open class words observed for

pauses in this experiment may in fact reflect only the previously observed difference between weak and strong syllables.

To test this hypothesis, we listened to the subjects' productions and noted whether the critical words had been produced with reduced or with full vowel quality. We found that for the two items *just* and *for* there was indeed significant vowel reduction in the baseline utterances; some subjects even reduced *for* in both repeated utterances as well as in the baseline. It is therefore possible that what appears to be a difference between closed and open class words is in fact the same difference between weak and strong syllables which we observed in the previous experiments.

6. General discussion

Our series of experiments allow several clear conclusions. Firstly, we have shown that speakers do mark word boundaries when they are trying to produce clear speech for a listener's benefit, and they do use durational manipulations to do so. The best evidence for this comes from the analysis of pausing in Experiments 1 and 2. When speakers had no knowledge of a need to speak clearly in the baseline utterances, they virtually never paused at a word boundary; but in deliberately clear repetitions they paused before words for as much as 150 ms or more. This suggests that speakers are aware that segmentation of continuous speech into words is a major problem for the listener.

Secondly, we have shown that speakers differentiate between boundaries which precede strong and weak syllables: on both the measures we examined, boundaries preceding weak syllables were in general marked more clearly than boundaries preceding strong syllables. The pattern of results in Experiments 3 and 4 confirm this interpretation. Word class cannot be a critical factor, since we did find differences in Experiment 3 in which word class was controlled, and we did not find differences systematically associated with word class in Experiment 4. Predictability cannot be a critical factor, since predictability was matched in Experiments 3 and 4, but

differences were nonetheless observed. Frequency cannot be a critical factor, since greater marking was applied to the higher frequency items in Experiments 1 and 2 but to the lower frequency items in Experiment 3.

Given that perceptual evidence has shown that listeners (to English) adopt a strategy of segmenting speech at strong syllable onsets, how should we interpret this result? The absence of frequency effects such as Cooper and Paccia-Cooper (1980) found for natural speech suggests that our effect is not simply an elaboration of natural speech effects, but is a strategy specifically associated with deliberately clear speech. Consistent with this, of course, is the absence of effects in our baseline utterances when speakers were unaware of any need to speak clearly. We suggest that in deliberately clear speech speakers are able to tailor their speech very efficiently to their listeners' needs. One way in which they do this is by compensating for the strategies which listeners normally use in speech perception. The English segmentation strategy of assuming that strong syllables are word-initial would fail to pick up boundaries before weak syllables; therefore to compensate for this strategy it is necessary to enhance the clarity particularly of boundaries preceding weak syllables. This is exactly what we have observed speakers to do in deliberately clear speech; by marking those word boundaries which would *not* be identified by application of the customary strategies, speakers are providing precisely the extra information which listeners most need.

Appendix: Experimental materials

For each experiment the six pairs of target sentences are listed, with the critical boundaries

marked "/", and the members of each pair labelled respectively S and W (for strong and weak) in Experiments 1 to 3, and O and C (for open and closed) in Experiment 4. The sentences in brackets are the first and second feedback sentences, in which the critical boundary is absent. The number following each target sentence is the predictability rating, i.e. the proportion (of 40 subjects) supplying the exact critical word in the cloze test.

Experiments 1 and 2

1S. Take it in / turns to eat breakfast (0.5)
(Baker interns all the terrorists; Take it internally at breakfast)
1W. He called in / to view it himself (0.975)
(The cold interviewer was selfish; He crawled into view by himself)

2S. It's colour of / eyes that counts (0.125)
(A skull will revise the count; His colours revise the counts)
2W. The waiter should serve / us first (0.125)
(Creating a nervous curse; The waiters put service first)

3S. The scene must be / headed "night" (0)
(A queen was beheaded tonight; The scene is beheaded knights)
3W. She tried to free / her trapped dog (0.025)
(The side of a rear trapdoor; She tried a freer trapdoor)

4S. A rich woman in / furs met us (0.5)
(Which one can infer some method; A rich woman infers matters)
4W. Get a yellow pill / for Mummy (0.65)
(Little fellows pilfer money; Better yell and pilfer money)

5S. Park in a nearer / place next year (0.15)
(How can an ear replace six ears; A car can near replace next year)
5W. The two / police entered the flat (0)
(A super leaf sent to attract; The super lease rented the flat)

6S. There's an / even split in the party (0)
(This uneven kitten departed; It's uneven splitting the party)

6W. The berry soup could poison / us too (0.175)
(This very super poisonous tool; The berry soup is poisonous too)

Experiment 3

1S. Play this / card a good deal more (0)
(They discarded woody board; They discard a good deal more)
1W. Fire this / cadet's automatic (0)
(Try and whisper bets all to Maggie; By a whisker debts ought to matter)

2S. We should view these paintings in / different light (0)
(The smoothie's pain is indifferent tonight; We should rue these paintings' indifferent light)
2W. He speaks of the fallen in / discreet terms (0)
(The speakers are calling indiscreet words; He speaks of the fall in indiscreet terms)

3S. Some big / insects bite nastily (0.1)
(The sum begins six by ten asked he; Some begin sects quite nastily)
3W. The ailing / inspector frightened his staff (0)
(The aliens pecked us and frightened us daft; The alien spectre frightened the staff)

4S. The new boss is / excellent and patient too (0)
(Two nurses select all the patients too; The new boss selects celebrations too)
4W. That chocolate is / expensive but nice (0.075)
(A lot of it isn't sensibly priced; That chocolate isn't sensible but nice)

5S. Why can't every / member pay a fee (0.05)
(I can't ever remember paying, did he?; Why can't Heather remember to pay a fee)
5W. A bust or / memorial is planned (0)
(The customer always plans; A customer more or less planned)

6S. He mended the / part for nothing (0)
(Remember departure puffing; The men did depart for nothing)
6W. Attach the / partition to the backing (0)
(That cheaper tissue will do for packing; Attach cheaper tissue to the packing)

Experiment 4

1O. Lots of / hour-long sessions are needed (0.025)
(Dogs devour long sections of needles; Lots avow a long session is needed)
1C. Both of / our children like peanuts (0.275)
(Don't devour children's ice cream much; Both devour children like peanuts)

2O. The ombudsman is a /just official (0)
(The comrades cannot adjust the fishes; The ombudsman is adjusting fishes)
2C. The new employees are / just fine (0.2)
(Renewal lawyers adjust fines; The new employees adjust fine)

3O. It heats to over / four hundred degrees (0)
(The feast won't afford under-threes; To heat stove, afford a hundred degrees)
3C. The work is over / for this year (0.625)
(The working cove affords a jeer; The work can so afford this year)

4O. It's hard to / buy decent clothes cheaply (0.675)
(Its art divides even close people; She can't abide recent clothes chiefly)
4C. The town is duller / by six o'clock (0.05)
(The townie's alibi fits a lot; The time for lullaby's six o'clock)

5O. Hold it on with a tighter / knot (0.1)
(Old as dawn is the Hottentot; Hold it on for a tie cannot)
5C. She wondered whether to dash or / not (0.275)
(Show underwear that Dad forgot; She wondered whether to mash shallots)

6O. It's colour of / eye that counts (0.125)
(A sculler arrived and pounced; His colour revived accounts)
6C. They wondered if / I was making the rules (0.075)
(To under-divide is breaking the rules; They want to defy his making the rules)

References

S. Butterfield and A. Cutler (1988), "Segmentation errors by human listeners: Evidence for a prosodic segmentation strategy", *Proceedings of SPEECH '88* (Seventh Conference of the Federation of Acoustic Societies of Europe), Edinburgh, Vol. 3, pp. 827-833.

F.R. Chen, V.W. Zue, M.A. Picheny, N.I. Durlach and L.D. Braida (1983), "Speaking clearly: Acoustic characteristics and intelligibility of stop consonants", *Speech Communication Group, MIT: Working Papers,* Vol. 2, pp. 1-8.

J. Clark, J. Lubker and S. Hunnicutt (1988), "Some preliminary evidence for phonetic adjustment strategies in communication difficulty", in *Language Topics: Essays in Honour of Michael Halliday,* ed. by R. Steele and T. Threadgold (John Benjamins, Amsterdam), pp. 161-180.

W.E. Cooper, J.M. Paccia and S.G. Lapointe (1978), "Hierarchical coding in speech timing", *Cognitive Psychology,* Vol. 10, pp. 154-177.

W.E. Cooper and J.M. Paccia-Cooper (1980), *Syntax and Speech.* (Harvard Univ. Press, Cambridge, MA).

A. Cutler and D. Carter (1987), "The predominance of strong initial syllables in the English vocabulary", *Computer Speech and Language,* Vol. 2, pp. 133-142.

A. Cutler and D.G. Norris (1988), "The role of strong syllables in segmentation for lexical access", */. of Experimental Psychology: Human Perception & Performance,* Vol. 14, pp. 113-121.

W.N. Francis and H. Kufera (1982), *Frequency Analysis of English Usage* (Houghton Mifflin, Boston).

F. Grosjean (1980), "Linguistic structures and performance structures: Studies in pause distribution", in *Temporal Variables in Speech,* ed. by H.W. Dechert and M. Raupach (Mouton, The Hague), pp. 91-106.

M.A. Picheny, N.I. Durlach and L.D. Braida (1986), "Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech", *J. of Speech and Hearing Research,* Vol. 29, pp. 434-446.

V.V. Valian and R.J. Wales (1976), "What's what: Talkers help listeners hear and understand by clarifying syntactic relations". *Cognition,* Vol. 4, pp. 115-176.