

Word boundary cues in clear speech: A supplementary report

Anne Cutler and Sally Butterfield

MRC Applied Psychology Unit, Cambridge, UK

Received 30 April 1991

Revised 7 August 1991

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. In an earlier report we showed that speakers do indeed mark word boundaries in clear speech, by pausing at the boundary and lengthening pre-boundary syllables; moreover, these effects are applied particularly to boundaries preceding weak syllables. In English, listeners use segmentation procedures which make word boundaries before strong syllables easier to perceive; thus marking word boundaries before weak syllables in clear speech will make clear precisely those boundaries which are otherwise hard to perceive. The present report presents supplementary data, namely prosodic analyses of the syllable following a critical word boundary. More lengthening and greater increases in intensity were applied in clear speech to weak syllables than to strong. Mean *F0* was also increased to a greater extent on weak syllables than on strong. Pitch movement, however, increased to a greater extent on strong syllables than on weak. The effects were, however, very small in comparison to the durational effects we observed earlier for syllables preceding the boundary and for pauses at the boundary.

Zusammenfassung. Eine der Hauptaufgaben für den Hörer beim Verstehen fließender Sprache ist die Zerlegung des Sprachsignals in Wörter. Wenn die Hörbedingungen ungünstig sind, kann der Sprecher dem Zuhörer durch bewusst deutliches Sprechen helfen. In vier Experimenten haben wir untersucht, wie Wortgrenzen in bewusst deutlich gesprochener Sprache erzeugt werden. In einem früheren Bericht haben wir gezeigt, dass Sprecher bei klarer Sprechweise Wortgrenzen tatsächlich markieren, indem sie zwischen den Wörtern Pausen einlegen und Silben vor Wortgrenzen verlängern; weiter werden diese Änderungen besonders auf Wortgrenzen vor schwachen Silben angewendet. Englische Zuhörer verwenden heuristische Unterteilungsstrategien, die dazu führen, dass Wortgrenzen vor starken Silben am einfachsten wahrzunehmen sind: die Betonung von Wortgrenzen vor schwachen Silben bei deutlicher Sprechweise markiert deshalb gerade die Grenzen, die sonst schwer zu erkennen wären. Dieser Beitrag beschreibt nachträglich die prosodischen Analysen, die wir an den Silben nach den kritischen Wortgrenzen ausgeführt haben. In schwachen Silben waren sowohl die Silbenverlängerung als auch die Intensitätserhöhung grösser als bei starken Silben; auch die Durchschnittsgrundfrequenz stieg mehr in schwachen als in starken Silben an. Die Tonhöhenbewegung dagegen nahm mehr zu in starken als in schwachen Silben. Im Vergleich zu den Verlängerungen der Dauer, die wir früher in Silben vor Wortgrenzen sowie bei Pausen zwischen Wörtern festgestellt haben, waren diese Veränderungen aber sehr klein.

Résumé. L'une des étapes les plus importantes de la compréhension du langage parlé est la segmentation du signal de parole en mots. Quand les conditions d'écoute sont mauvaises, les locuteurs peuvent aider ceux qui les écoutent en articulant bien, c'est-à-dire en produisant une "parole claire". Nous avons examiné, au cours de quatre expériences, la production des frontières de mots dans la parole claire. Un rapport précédent a montré que les locuteurs qui articulent très soigneusement essaient effectivement de souligner les frontières de mots en effectuant une pause aux frontières et en allongeant les syllabes qui les précèdent. De plus, ils appliquent particulièrement ces stratégies aux frontières qui précèdent les syllabes faibles. Les anglophones emploient en effet une stratégie de segmentation telle qu'ils repèrent plus aisément les frontières de mots précédant les syllabes fortes ou accentuées; le soulignement des frontières de mots précédant les syllabes faibles dans la parole claire permet de marquer les frontières qui seraient difficiles à percevoir. Le présent article décrit des données supplémentaires: les analyses prosodiques des syllabes qui suivent une frontière critique. Dans la parole claire, la durée est plus allongée et l'intensité davantage renforcée dans les syllabes faibles que dans les syllabes fortes. La fréquence fondamentale moyenne augmente aussi davantage dans les syllabes faibles que dans les syllabes fortes. Au contraire, le mouvement du ton augmente plus dans les syllabes fortes que dans les syllabes faibles. Ces effets sont toutefois très petits comparés aux allongements, observés auparavant, de la durée des syllabes précédant les frontières de mots et de celle des pauses à la frontière.

Keywords. Speech production, intelligibility, clear speech, word boundaries, segmentation, intonation, pitch, loudness, intensity.

1. Introduction

Continuous speech presents the listener with a lexical segmentation problem: to understand the speech it is necessary to identify within it the individual tokens which correspond to entries in the listener's lexicon – but where does one such item finish and the next begin? Except at each end of a complete utterance, robust and reliable cues to word boundaries hardly exist.

Studies of perceptual errors show that listeners quite often make mistakes with lexical segmentation. Bond and Garnes (1980), Browman (1978) and Cutler and Butterfield (1992), for example, describe many errors of this type: "by loose analogy" is perceived as "by Luce and Allergy", "a must to avoid" is perceived as "a muscular boy", and so on.

It is a reasonable assumption, therefore, that one way speakers could really help listeners would be to remove the lexical segmentation problem entirely, by making word boundaries explicit. Such explicit segmentation of most normal utterances would, we assume, clash with the demands of rhythmic continuity in both speech production and perception. However, situations do arise in which speakers deliberately try to speak clearly because they have become aware that listeners are having trouble. In such situations, explicit word boundary cues would presumably prove an efficient communicative aid.

The project on which we report here was prompted by the discovery that previous studies of deliberately clear speech (e.g. (Chen et al., 1983; Clark et al., 1988; Paul-Brown and Yeni-Komshian, 1988; Picheny et al., 1986)) had paid relatively little attention to word boundaries in continuous utterances. Picheny et al. (1986) reported that clear speech contains pauses at word boundaries, although most such pauses were much shorter than the 250 msec which is commonly used as the threshold for defining a pause in other studies of speech production (Grosjean, 1980). Certainly, no previous study had compared explicit segmentation cues across different *types* of word boundary.

In four experiments we investigated this question; the present paper reports one aspect of our results, namely the prosodic characteristics of syl-

lables which follow a word boundary (i.e. word-initial syllables). To show how this report fits into our study as a whole, we begin with a summary of the clear speech corpus which we collected for the project.

1.1. *The corpus*

In the present study we collected a corpus of utterances which were spoken deliberately clearly, and compared them with the same sentences uttered by the same speakers without any attempt to be especially clear. We compiled three separate sets of twelve sentences, and each set was recorded by ten speakers. Each speaker recorded each sentence three times – once without trying to be clear, and twice more under the impression that a listener was having difficulty hearing; we assumed that these latter two utterances would be designed to be clear.

These manipulations produced a total of ($3 \times 12 \times 10 \times 3 =$) 1080 utterances. Readers will appreciate that this is an enormous corpus of speech to analyse. We have, therefore, rendered the task manageable by restricting our analysis in two major ways¹. Firstly, we confined our attention to the region of one crucial word boundary in each utterance. We defined this boundary region as beginning at the onset of the syllable preceding the boundary and ending at the offset of the syllable following the boundary. Secondly, we investigated only the prosodic characteristics of the utterances. (At least one previous study, by Clark et al. (1988) found a greater amount of change in suprasegmental than in segmental measures of clear versus baseline speech.)

Even with these restrictions, the corpus offers a very large body of data. The domain of analysis covers two syllables in each utterance, plus, of course, any between-syllable interval. The dimensions of prosodic structure include syllable duration, intensity and fundamental frequency.

We measured, therefore, the duration of the pre-boundary syllable; the duration of the between-syllable interval (pause at the boundary), if any; the duration of the syllable following the

¹ We are happy to make the corpus available to researchers who would be interested in exploiting it further.

boundary; the mean intensity of the pre-boundary syllable; the mean intensity of the post-boundary syllable; the mean fundamental frequency (F_0) of the syllable preceding the boundary, and the standard deviation of F_0 ; and, finally, the mean F_0 of the syllable following the boundary, and again, the standard deviation. For each of these nine sets of measurements, we conducted two separate analyses of variance, one with subjects as random factor and one with sentences as random factor, accepting as reliable only differences which were statistically significant in both analyses.

Two of these sets of measurements have already been published. In Cutler and Butterfield (1990) we reported the durational measures of pre-boundary syllables and between-syllable pauses. The present report describes a further four sets of measurements, namely the four (duration, intensity, mean F_0 , F_0 standard deviation) conducted on the word-initial, i.e. post-boundary, syllables.

The remaining three sets of measurements (intensity, mean F_0 and F_0 standard deviation of the pre-boundary syllables) did not yield results of interest and will not be reported in detail here².

1.2. *Independent variables*

The word boundaries which we examined varied in certain systematic ways. The principal factor which we manipulated was whether the syllable following the boundary was strong (i.e., contained a full vowel) or weak (i.e., contained a reduced vowel). This factor has been shown to be important in the recognition of English speech. Most lexical words in English begin with strong syllables (Cutler and Carter, 1987), and listeners appear to be able to exploit this fact as a way of solving the problems inflicted upon them by the absence of reliable word boundary cues. Laboratory studies have suggested that listeners segment speech signals at the onset of strong syllables (Cutler and Norris, 1988), and studies of misperceptions show a pattern consistent with such

behaviour: both in spontaneous slips and in laboratory-induced misperceptions, listeners tend to erroneously insert word boundaries more often before strong than before weak syllables (Cutler and Butterfield, 1992).

In three of our four experiments, therefore, half of the critical post-boundary syllables were strong and the other half were weak. In English there is a significant tendency, however, for syllable strength to vary with word class – open-class (lexical, or content) words tend to begin with strong syllables, but closed-class (grammatical, or function) words are mostly monosyllabic and weak in continuous speech. In the fourth experiment, therefore, we conducted a separate comparison test of the effect of word class. This was done in the same manner in which we had varied syllable strength, so that in Experiment 4 half the post-boundary syllables were open-class words and the other half were closed-class words.

We also assessed the role of word frequency and of predictability in determining the structure of clear speech. Cooper and Paccia-Cooper (1980) studied certain phonological rules which can apply across word boundaries, and have the effect of obscuring word onsets (e.g. palatalisation, which turns *did you* into [dɪdʒu] and effectively obscures the beginning of the word *you*). They reported that speakers are less likely to apply such rules across a word boundary if the word after the boundary has a low frequency of occurrence than if it is of high frequency. A possible interpretation of this result is that low-frequency words are less predictable than high-frequency words, and speakers are reluctant to make unpredictable words harder for listeners to hear. We tested this effect in our own study by determining the frequency of occurrence of the post-boundary word in all our sets of materials. In one of the two sets which varied syllable strength, weak post-boundary syllables had higher frequency than strong, while in the other set the strong post-boundary syllables had higher frequency than the weak. We assessed the effect of predictability separately by determining how easily a group of native speakers could supply the post-boundary word given its context: in two of our sets of materials, predictability was matched across the principal factor under investigation.

² Details of these unpublished analyses are available from the authors on request. A partial report of analyses conducted on the F_0 measures of the pre- (as well as the post-)boundary syllables may be found in Butterfield and Cutler (1990).

The main objective of our project was to see whether speakers do mark boundaries at all in clear speech. If they do, we then wanted to know whether they mark some types of word boundary more than others: boundaries before strong or before weak syllables; boundaries before open-class or before closed-class words; boundaries before high-frequency or before low-frequency words; boundaries before relatively predictable or before relatively unpredictable words.

The durational measurements of the pre-boundary syllable and of inter-syllable intervals which we have already reported (Cutler and Butterfield, 1990) showed that speakers do indeed mark word boundaries in clear speech, by lengthening syllables before the boundary, and by pausing at the boundary. Frequency, predictability and word class of the post-boundary word had no systematic effects on these measures, but there was a significant effect of the syllable strength factor – speakers paid significantly more attention to marking boundaries before weak syllables than before strong syllables. That is, they tend to mark most strongly the very boundaries which listeners would *not* detect by application of the normal (strong-syllable-based) segmentation procedures. We argued in our previous report that this pattern indicates that speakers' clear speech strategies are particularly well attuned to listeners' needs.

In the present report we describe how speakers produce word-initial, i.e. post-boundary, syllables in deliberately clear speech.

2. The experiments

2.1. Method

2.1.1. Materials

The materials for Experiments 1 and 2 were twelve sentences of relatively unpredictable content. 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 preceding and following 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 designed to exploit possible mishearings in which the critical boundary was absent. For instance, "in turns" could be misheard as "interns", while "in to view" could be misheard as "interview". 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".

As described more fully in Cutler and Butterfield (1990), this first set of materials used in Experiments 1 and 2 did not control some potentially relevant variables. For instance, the word frequency was quite strongly mismatched because of a word class mismatch: five of the six weak syllables were high frequency closed class words. Thus the frequency of occurrence (determined from the Francis and Kučera (1982) word-class-specific norms) for the weak post-boundary syllables was in general much higher than that of the strong syllables. In the two further experiments we used materials in which both word class and frequency were much more closely controlled. 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. In particular, Experiment 3 offers a more closely controlled investigation of the strong-weak com-

parison addressed in Experiments 1 and 2.

For Experiment 3 a second 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, in four of six pairs, 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. Purported mishearings were again constructed for use as feedback.

A third set of 12 sentences was constructed for Experiment 4, again in six matched pairs. In this case the critical variable was word class of the post-boundary word; homophones were chosen which could be either open- 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". Each closed class word was higher in frequency than its open class pair. 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 chose contexts in which reduction was unlikely. Purported mishearings were again constructed for use as feedback.

The predictability of the post-boundary words in all three materials sets was assessed by means of a cloze text. Forty 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. In the first materials set, 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. In the second materials set, predictability was matched in that all items received scores of virtually zero predictability. In the third materials set, 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.

All three sets of target and feedback sentences, with the predictability ratings for the post-boundary words in each case, are listed in Appendix A.

2.1.2. Subjects

Five members of the Applied Psychology Unit subject panel, four females and one male, took part (for payment) in Experiment 1. A further five, three females and two males, took part in Experiment 2. Ten subjects from the same population, seven females and three males, took part in Experiments 3 and 4, and again were paid for their participation. All subjects were native speakers of British English.

2.1.3. Procedure

In Experiment 1, the subjects 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; this response would be displayed on the subjects' VDU. 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.

Although these instructions specifically re-

requested that each sentence be spoken naturally on its first rendition, the possibility exists that awareness of the possibility of listener difficulty, coupled with carryover effects from clear speech required in repetitions on previous trials, might result in even first renditions having the characteristics of clear speech. We therefore modified the instructions to subjects. Thus in Experiment 2, the subjects produced the same sentences as in Experiment 1, under the same conditions, except that they read all the experimental (and filler) sentences aloud onto tape right at the beginning of the experimental session, before they were told about the listener and the supposed distortion. These initial productions then served as the baseline to be compared with the two post-feedback repetitions.

This latter procedure was also used in Experiments 3 and 4. These two experiments were administered together, so that the total number of items, including the three practice and ten filler sentences, was in this case 37.

For each subject in each experiment, the baseline production and both repetitions of each experimental sentence were digitised at a sampling rate of 10 kHz. The syllable before and after each critical boundary was marked and measured, using the Camsed speech editor. Syllable durations were recorded in milliseconds.

For the pitch analysis, we used the Schafer-Vincent (1983) algorithm; this algorithm determines, in several stages, the most likely location of glottal pulses in the voiced (i.e. quasi-periodic) portions of the waveform. The stages include analysis of the periodic structure of an amplitude-against-time representation of the waveform; the reciprocal of the distance between successive pulses is used as an instantaneous measure of the frequency of glottal pulse occurrence, i.e. $F0$. The output of the algorithm, therefore, is a sequence of $F0$ values, one value corresponding to each glottal pulse. Thus the number of $F0$ values calculated for any syllable varies according to the length of the syllable and the $F0$ of the voiced portions. Within each syllable, the mean and standard deviation of the $F0$ values were calculated. The mean $F0$ value across each syllable gives an estimate of relative pitch *height*; the standard de-

viation of this mean for each syllable gives an estimate of relative pitch *movement* within the syllable. We analysed each measure separately.

The algorithm failed to calculate values for a small proportion of the syllables in each experiment – in total, 146 post-boundary syllables (i.e. 13.5% of the 1080) failed to receive a value. After inspecting the algorithm's output, we discarded a further 16 syllables (1.5%) as possible aliasing errors. In these cases the missing data point was replaced by the subject's average for that condition.

We also measured the intensity of the syllables before and after the critical boundary. Although the subjects had been instructed not to shout, we observed some increase of speech intensity in the clear utterances compared to the baseline. For the intensity measurements, we used an algorithm which calculated the root mean square amplitude for each analysed syllable, and expressed the values on a decibel scale on which 0 dB corresponds to a sine wave with amplitude equal to one minimum quantum of input voltage.

As Cutler and Butterfield (1990) pointed out, it is inappropriate to conduct a statistical analysis on absolute measurement values where the syllables to be compared have considerable intrinsic differences (for instance in number of phonemes) which might be expected to affect the dimensions of measurement. This is particularly relevant in the case of the post-boundary syllables, because the main independent variable is the nature of these syllables. It is to be expected that strong and weak syllables, for instance, will exhibit intrinsic differences on all our measurement dimensions: duration, $F0$, intensity. Accordingly, we conducted our statistical analyses not on the absolute values but on the *proportional increase* in each repetition in comparison with the baseline. The figures, however, show the absolute values (from which the reader will readily be able to confirm that intrinsic differences – e.g., between strong and weak syllable duration – do indeed exist).

Note that the proportional increase from baseline to repetitions shown by the absolute means (in the figures) will not necessarily correspond exactly to the means of the individual proportional increases for each subject on each sen-

tence (on which the analysis of variance was conducted)³. We report in the text the means of the proportional increases.

Two analyses of variance were conducted on the measures of proportional increase on each dimension, one with subjects and one with sentences as random factor. We report only main effects (or interactions) which reached significance in both analyses, i.e. can with some confidence be generalised to the wider population of speakers and sentences.

3. Experiments 1 and 2

The materials for these two experiments (listed in full in the Appendix) compared boundaries before strong versus before weak syllables. The post-boundary words were: *turns, furs, eyes, headed, place, even* (strong) and *to, for, us, her, police, us* (weak). As can be seen, most of these are monosyllabic, i.e. most of the post-boundary syllables were themselves words.

The strong syllables will in general be longer and louder than the weak syllables, and will exhibit a greater amount of pitch movement. These intrinsic differences should be present in the baseline utterances. We may expect increase on all measures in clear speech compared with the baseline. Moreover, since the weak syllables are mostly monosyllabic words, we might expect that

they may in some cases be turned into strong syllables, by removing vowel reduction; for instance, *to* in *in to view* would normally be pronounced [tə], but in deliberately clear speech it might be pronounced [tu]. The effect of this could be to bring the measures for the weak syllables on at least some dimensions closer to those for the strong syllables.

In order to determine whether this did occur, both authors and a phonetically trained colleague listened to all the subjects' utterances and transcribed the vowels in the post-boundary syllables. In the weak-syllable condition of Experiment 1 we found that 20.7% of first repetitions and 27.6% of second repetitions contained full vowels which had been reduced in the baseline; in Experiment 2 the figures were 20% and 36.7%, respectively. Thus turning a reduced vowel into a full vowel in clear speech was by no means the rule, but it did occur on a potentially significant number of occasions. (In Experiment 1 48.3% of first repetitions and 41.4% of second repetitions contained reduced vowels unchanged from the baseline, while 27.6% of both repetitions contained full vowels unchanged from the baseline; in one case a subject produced a full vowel in the baseline but a reduced vowel in each repetition. In Experiment 2 40% of first repetitions and 23.3% of second repetitions contained reduced vowels unchanged from the baseline, while 40% of both repetitions contained full vowels unchanged from the baseline.)

3.1. Duration

Figures 1(a) and 1(b) show the mean durations (across subjects and sentences) of the post-boundary syllables in Experiments 1 and 2, respectively. In Experiment 1, the first and second repetitions of the strong syllables were on average 13.8% and 18.3% longer than the baseline; for the weak syllables, the mean proportional increases were 23.2% and 40.5%, respectively. In Experiment 2, the strong syllables increased by 47.9% and 57.3%, respectively, while the weak syllables increased by 75.5% and 96.7%, respectively. Thus the proportional increases were consistently greater for weak syllables than for strong. However, analysis of variance (of these

³ The differences will usually be trivial, except where there is missing data and high variability in the data set. The missing data condition is met in our analyses of *FD*. Here, discrepancies were particularly noticeable in the standard deviations, where the variability was also very high, in that a few subjects on a few words markedly altered the amount of pitch movement they produced, but such behaviours were inconsistent and infrequent. Exacerbating the problem is the fact that the values being reported are very small (typically 10 Hz or less). In Experiment 3, for example, the standard deviation for strong syllables in the baseline is 6.7 Hz, and in the first repetition 12.3 Hz. The second number is 84% higher than the first; but the mean of the individual proportional increases is 125%. Since, of course, the high variance is itself the major result of the measurement we report here, and is reflected in the lack of significance in these analyses, these discrepancies are of little practical importance. However, for the sake of scrupulous completeness we have provided both sets of figures for the reader: one in the figures, the other in the text.

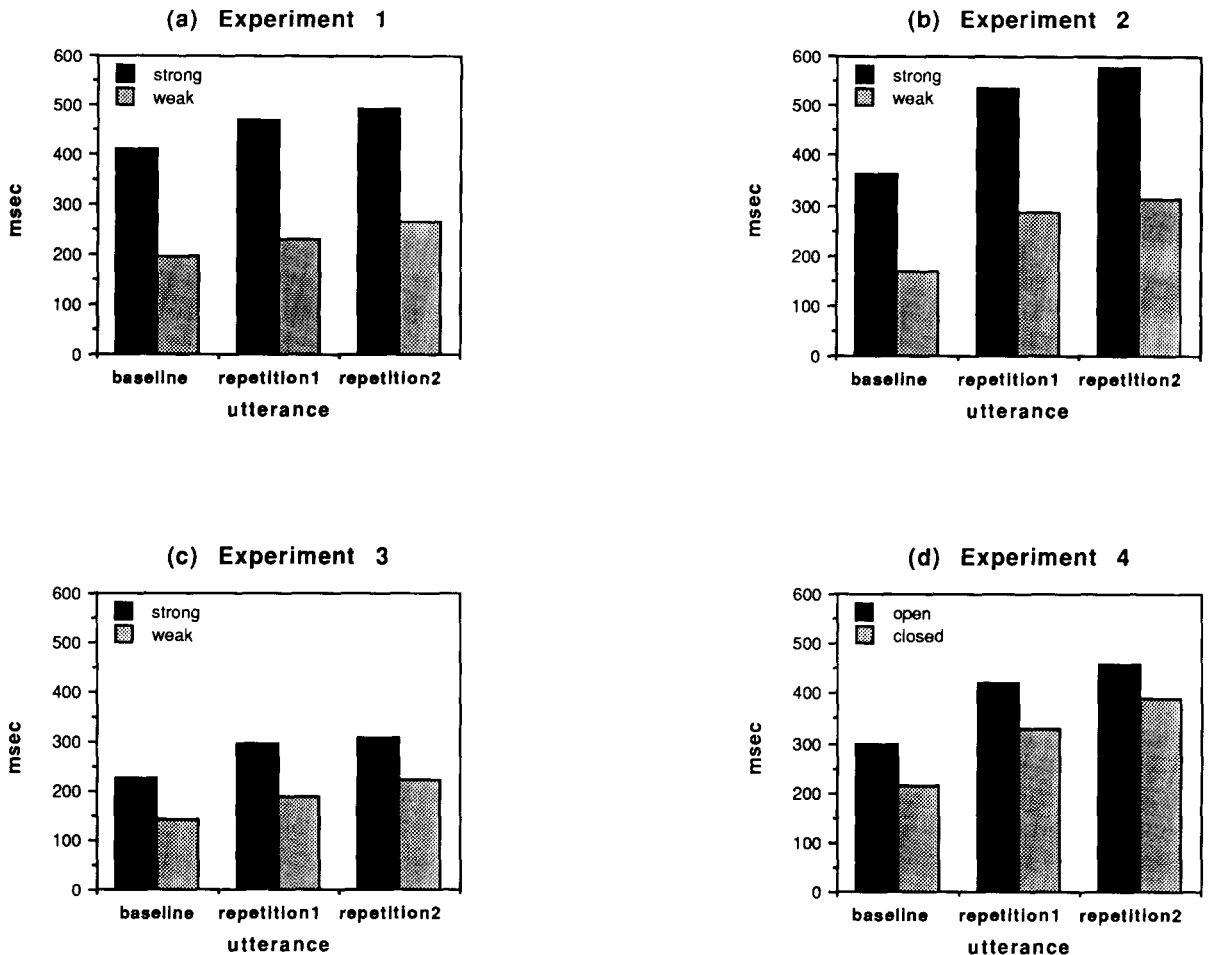


Fig. 1. Mean duration (msec) of the syllable following the boundary in the baseline, first repetition and second repetition, separately for strong and weak syllables in (a) Experiment 1, (b) Experiment 2 and (c) Experiment 3, and for open and closed class words in (d) Experiment 4.

proportional increases) revealed no significant differences between strong and weak syllables in either experiment.

3.2. Pitch

Figures 2(a), 2(b), 3(a) and 3(b) show the means (across subjects and sentences) of the means and standard deviations of F_0 of the post-boundary syllables in Experiments 1 and 2. The pitch algorithm failed to compute values for 10.8% of all the syllables in these experiments. The values computed for a further 1.9% of post-boundary syllables were discarded as probable aliasing errors.

In Experiment 1, mean F_0 of both strong and weak syllables showed a slight *decrease* in repetitions in comparison to the baseline. For strong syllables, the mean decrease was 1.9% and 3.6% for the first and second repetition, respectively; for weak syllables, the corresponding values were 0.9% and 1.5%. In Experiment 2, there were equally small increases: 2.2% and 0.4%, respectively for strong syllables, 3.6% in each case for weak syllables. In neither case was there a significant difference between the two syllable types.

The analysis of standard deviations in Experiment 1 showed that for both strong and weak syllables there was more pitch movement in repetitions than in the baseline: the proportional in-

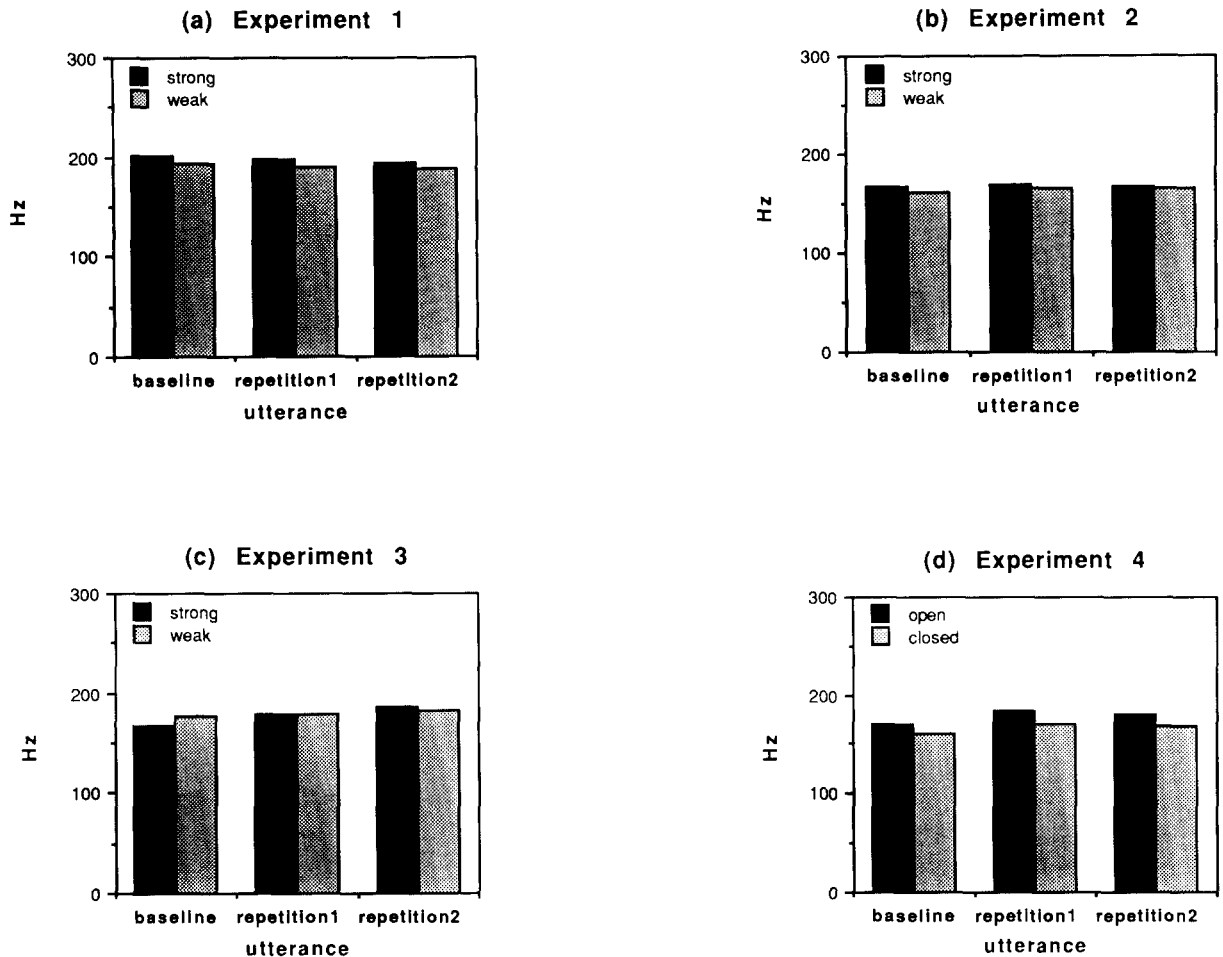


Fig. 2. Mean F_0 (Hz) of the syllable following the boundary in the baseline, first repetition and second repetition, separately for strong and weak syllables in (a) Experiment 1, (b) Experiment 2 and (c) Experiment 3, and for open and closed class words in (d) Experiment 4.

creases were 39.4% and 51.3%, respectively, for strong syllables, 9.5% and 15.3%, respectively, for weak syllables. In Experiment 2 the pitch movement on strong syllables increased on average by 110% and 101%, respectively, while on weak syllables the amount of movement increased by 31.5% on the first repetition but by only 5.6% on the second. The analysis of variance revealed no statistically significant effects in Experiment 1, but in Experiment 2 the advantage of strong over weak syllables almost reached significance: $F_1 [1,4] = 6.44, p < 0.07$; $F_2 [1,10] = 17.12, p < 0.01$.

3.3. Intensity

Figures 4(a) and 4(b) show the mean intensity (across subjects and sentences) of the post-boundary syllables in Experiment 1 and 2. In Experiment 1 there was a slight *decrease* in intensity of strong syllables in repetitions compared with the baseline, by 3.1% and 1.9%, respectively, but an equally slight increase in intensity of weak syllables, by 1.1% and 2.4%, respectively. In Experiment 2 there was virtually no change in strong syllables (0.2% and 0.6% increase, respectively) but small increases (6.8% and 8.9%) for weak syllables. In neither experiment did these differ-

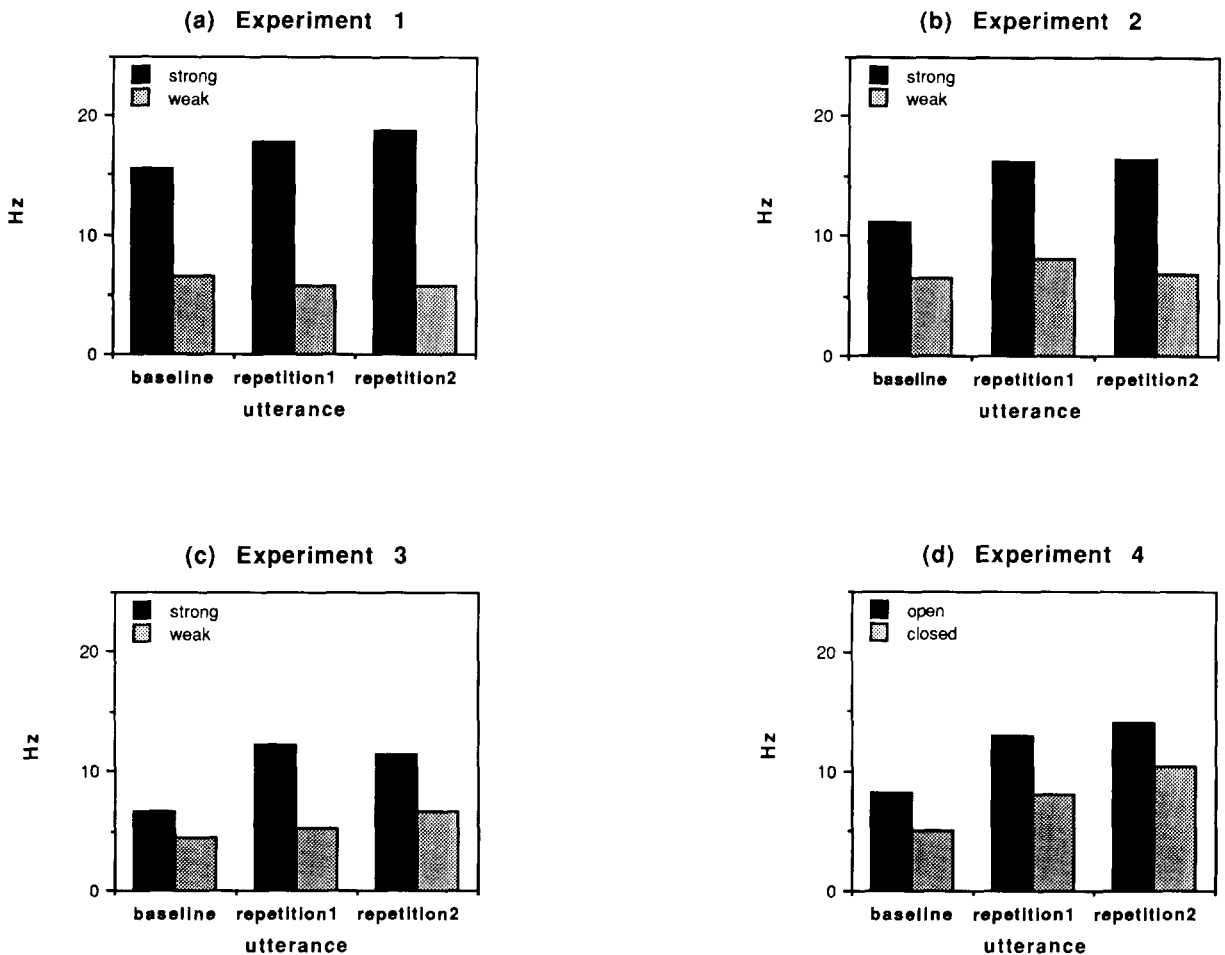


Fig. 3. Mean of the standard deviations of the F_0 (Hz) of the syllable following the boundary in the baseline, first repetition and second repetition, separately for strong and weak syllables in (a) Experiment 1, (b) Experiment 2 and (c) Experiment 3, and for open and closed class words in (d) Experiment 4.

ences reach significance in the analysis of variance.

4. Experiment 3

The materials for Experiment 3 again compared boundaries before strong versus before weak syllables; the post-boundary words were *card*, *different*, *insects*, *excellent*, *member*, *part* (strong) and *cadet*, *discreet*, *inspector*, *expensive*, *memorial*, *partition* (weak). Nearly all of these words are polysyllabic. We would, therefore, expect that subjects would be much less likely to turn syllables which were weak in their baseline

productions into strong syllables in clear speech. Altering vowel quality significantly impairs word recognition (Bond and Small, 1983; Cutler and Clifton, 1984), and such alteration would be unlikely to act as a useful aid to listeners.

Indeed, phonetic transcription of the utterances from this experiment showed that in the weak-syllable condition only 5% of first repetitions and 15% of second repetitions contained a full vowel where the baseline utterance had contained a reduced vowel; 81.7% of first repetitions and 71.7% of second repetitions contained a reduced vowel unchanged from the baseline. (The remaining 13.3% of each repetition type contained a full vowel unchanged from the baseline.)

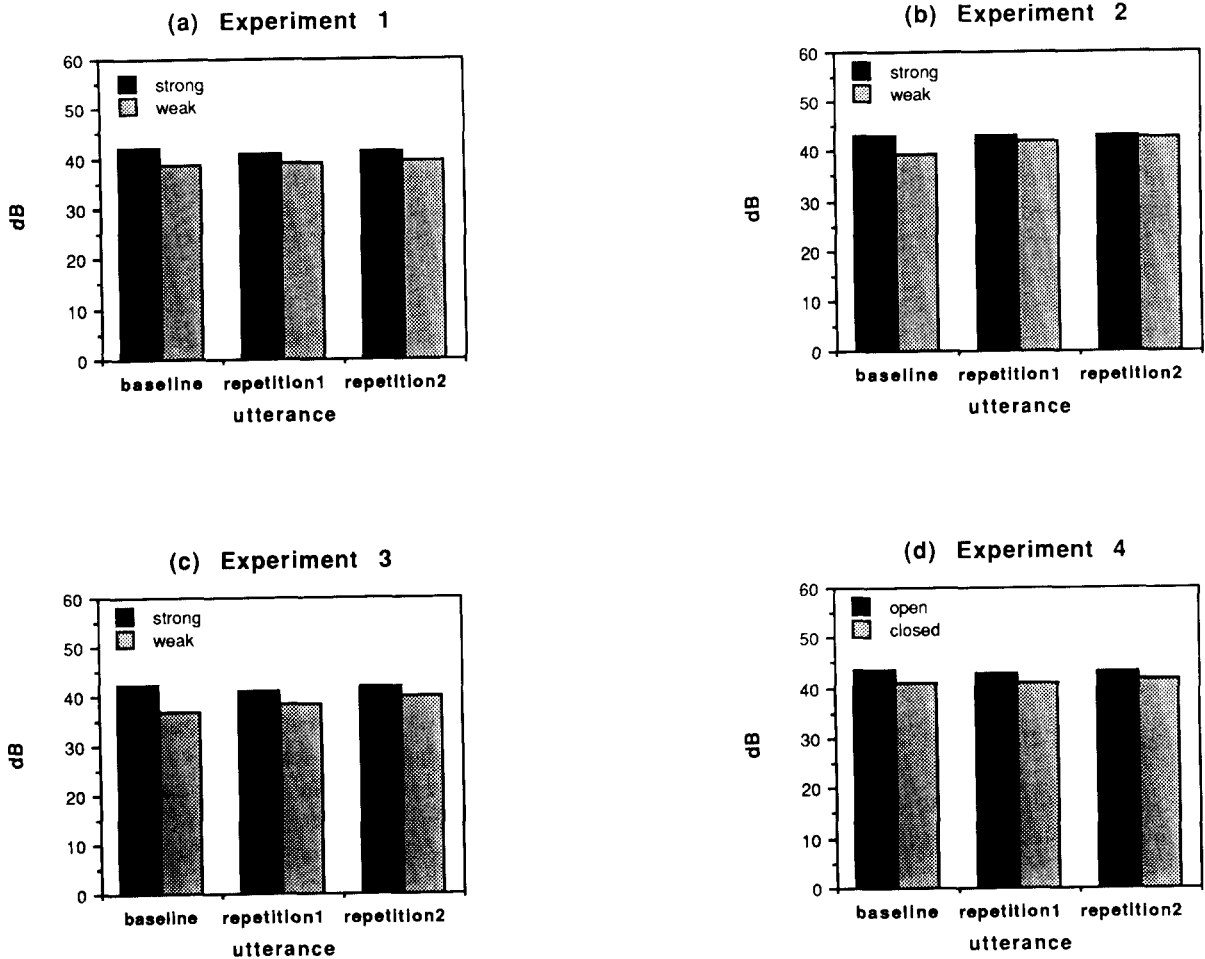


Fig. 4. Mean intensity (dB) of the syllable following the boundary in the baseline, first repetition and second repetition, separately for strong and weak syllables in (a) Experiment 1, (b) Experiment 2 and (c) Experiment 3, and for open and closed class words in (d) Experiment 4.

4.1. Duration

Figure 1(c) shows the mean durations (across subjects and sentences) of the post-boundary syllables in Experiment 3. In comparison to the baseline, durations of these syllables increased in the repetitions by 29.4% and 33.9%, respectively, for strong syllables, and by 35% and 63.6%, respectively, for weak syllables. Once again, the greater amount of increase in weak syllables as opposed to strong did not, however, result in a statistically significant effect in the analysis of variance.

4.2. Pitch

Figures 2(c) and 3(c) show the means (across subjects and sentences) of the means and standard deviations of F_0 of the post-boundary syllables in Experiment 3. Algorithm failures comprised 18.6% of all the syllables in this experiment; a further 0.8% of post-boundary syllable values were discarded as probable aliasing errors. The analysis of mean F_0 revealed again slightly higher F_0 in repeated than in baseline utterances, the proportional increases being 8.9% and 13.1%, respectively, for strong syllables, and 16.8% and 21%, respectively, for weak syllables. The differ-

ence between strong and weak syllables was again statistically insignificant. The amount of pitch movement also increased, by 125% and 106%, respectively, for strong syllables, and by 48.1% and 87.1%, respectively, for weak syllables. Again, however, these differences were not statistically significant.

4.3. Intensity

Figure 4(c) shows the mean intensity (across subjects and sentences) of the post-boundary syllables in Experiment 3. The mean intensity of strong syllables decreased slightly in repetitions in comparison to the baseline (by 2% and 0.2%, respectively), while the mean intensity of weak syllables increased (by 4.7% and 9%); in this case the greater amount of increase on weak syllables than on strong was reflected in statistical significance in the analysis of variance: $F_1 [1,10] = 10.38, p < 0.01, F_2 [1,10] = 21.72, p < 0.005$.

5. Experiment 4

The materials for this experiment compared boundaries before open-class versus closed-class words. The items were chosen because they were homophones, so that the post-boundary words were phonologically identical. The open-class set was *hour, just, four, buy, knot, eye*, while the closed-class set was *our, just, for, by, not, I*.

Although the closed-class set was chosen to be as far as possible resistant to vowel reduction, in some cases (as reported in Cutler and Butterfield (1990)) subjects did reduce the words *just, for* and *not* in the baseline. We found that 23.3% of first repetitions and 26.7% of second repetitions contained full vowels where the same speaker had produced a reduced vowel in the baseline utterance. (8.3% of first repetitions and 5% of second repetitions contained a reduced vowel unchanged from the baseline, while the remaining 68.3% of each repetition type contained a full vowel unchanged from the baseline.)

5.1. Duration

Figure 1(d) shows the mean durations (across subjects and sentences) of the post-boundary syl-

lables in Experiment 4. Durations were longer in the two repetitions than in baseline utterances, by 48.6% and 61.5%, respectively, for open-class words, and by 66.4% and 100.8%, respectively, for closed-class words. The analysis of variance showed no significant differences between open- and closed-class words.

5.2. Pitch

Figures 2(d) and 3(d) show the means (across subjects and sentences) of the means and standard deviations of F_0 of the post-boundary syllables in Experiment 4. Algorithm failures comprised 11% of all syllables, and a further 1.7% of post-boundary syllable values were discarded as probable aliasing errors. The analysis of mean F_0 showed higher F_0 on the two repetitions than on the baseline utterances, the proportional increases being 10.4% and 8.6%, respectively, for open-class words and 7.8% and 7.1%, respectively, for closed-class words. Again, there were no significant effects in the analysis of variance.

The analysis of standard deviations also showed increases in the two repetitions in comparison to the baseline: for open-class words the increases were 145% and 187%, respectively, while for closed-class words they were 167% and 215%, respectively. Yet again, there were no significant effects in the analysis of variance.

5.3. Intensity

Figure 4(d) shows the mean intensity (across subjects and sentences) of the post-boundary syllables in Experiment 4. The intensity of open-class words hardly differed in the repetitions compared to the baseline (a mean decrease of 0.04% occurred in the first repetition, and a mean increase of 0.02% in the second); for closed-class words there was a slight increase in intensity in repetitions (1.4% and 3.2%, respectively). The analysis of variance showed no effects reaching the required level of significance.

6. Summary

The durational measurements showed that in all of Experiments 1, 2 and 3 there was a greater

proportional lengthening of weak syllables than of strong syllables when speakers were aiming for clear speech. In Experiment 4 closed-class words were lengthened to a greater extent than open-class words. However, none of these effects were statistically significant.

Mean *F0* levels in clear speech likewise rose to a greater extent on weak than on strong syllables in Experiments 2 and 3, and dropped less on weak than on strong syllables in Experiment 1. In Experiment 4 open-class words showed a greater proportional rise in *F0* than closed-class words. Again, no effects at all reached statistical significance.

The standard deviations of *F0*, as a measure of pitch movement, revealed greater increases in movement for strong syllables than for weak syllables in Experiments 1, 2 and 3, but in only one case did the difference approach significance. In Experiment 4 there was, however, a greater (though, again, statistically insignificant) increase in movement on closed-class words.

Finally, intensity measurements showed that intensity increased in clear speech to a greater extent on weak than on strong syllables in all of Experiments 1, 2 and 3, and to a greater extent on closed-class words in Experiment 4. The difference in Experiment 3 was statistically reliable.

The overall picture, therefore, is that speakers who are trying to be particularly clear tend to lengthen syllables, raise pitch slightly, slightly increase intensity (even when asked not to shout), and increase pitch movement.

Although speakers manipulate the prosody of both strong and weak syllables in the same way, it is noteworthy that the effects tend to be rather greater on weak syllables. (Although very few differences reached significance, of the twelve strong/weak comparisons – four measures in each of Experiments 1, 2 and 3 – nine showed stronger effects in weak than in strong syllables.) Open-class words in general pattern like strong syllables, while closed-class words tend to pattern like weak syllables (three out of four comparisons showed stronger effects on closed- than on open-class words).

As we noted above, one way in which weak syllables can be made especially clear is by turning them into strong syllables. It is unlikely that the

significant effects which did appear resulted from vowel quality alterations of this kind, since our listening tests showed that such alterations occurred in only about a quarter of cases in Experiments 1, 2 and 4, and considerably less often in Experiment 3. However, in order to assess the effect of vowel quality alterations on the differences we observed, we undertook a correlational analysis. For every sentence with a critical weak syllable in every experiment, we computed the number of speakers who had altered vowel quality in each repetition in comparison with their baseline production. We then computed the correlation between these values and the mean proportional changes from baseline to each repetition for the same sentences, on each of the four measures. Thus there were 16 analyses in all, four for each of the four experiments. Exactly one correlation reached significance for each experiment, but there was no consistency in the pattern of effects. In Experiment 1, there was a significant correlation ($r [11] = 0.79, p < 0.01$) between the frequency of vowel quality alteration and the amount of durational increase. The same was true in Experiment 3 ($r [11] = 0.75, p < 0.01$). In Experiment 2 the frequency of vowel quality alteration was significantly correlated with the amount of increase in mean *F0* ($r [11] = 0.6, p < 0.05$). In Experiment 4 there was a significant correlation ($r [11] = 0.63, p < 0.03$) between the frequency of vowel quality alteration and the amount of increase in intensity. The twelve remaining correlations, including all correlations with amount of pitch movement, failed to reach significance. Since there was no clear pattern in these correlations (and, in particular, the correlations were not stronger where we had observed stronger effects in the analyses of variance), we conclude that vowel quality alteration is not the major factor producing prosodic differences between weak and strong syllables in clear speech.

The principal independent variable which we manipulated, namely the strong-weak syllable contrast, does therefore seem to have produced a small but consistent pattern of differences. At least for duration and intensity, effects are consistently greater on weak syllables than on strong; mean *F0* shows a tendency (though slighter) towards the same pattern; pitch movement tends to

increase to a greater extent on strong than on weak syllables. We did not, however, find systematic effects of word class, except where word class was fully confounded with the strong-weak contrast – that is, we found differences between the closed and open word classes in Experiment 4, where word class was manipulated but the strong-weak contrast not fully factored out, but we also found differences between strong and weak syllables in Experiment 3, in which word class was fully controlled. We also found no effects of word frequency – there were more prosodic effects on higher frequency words in Experiments 1, 2 and 4, but on lower frequency words in Experiment 3. Likewise, there were no effects of predictability – or rather, in Experiments 3 and 4 we *did* find differences between materials sets in which predictability was invariant.

7. General discussion

When speakers try to help listeners by speaking deliberately clearly, they usually do the right thing. The ways in which they alter their speech output result in signals which are indeed easier for listeners to interpret.

At the phonetic level, clear speech contains less ambiguity in cues to phonetic identity. Picheny et al. (1986) (see also Chen et al. (1983)) found that the VOT distributions for voiced versus unvoiced stop consonants in clear speech were fully separated, whereas in speech which is not deliberately clear, the same distributions tend to overlap (Chen et al., 1983; Klatt, 1975). Likewise, word-final stops tend to be fully released in clear speech; the release provides clear information about both voicing and place of articulation (Malécot, 1958). Again, this information is often missing in speech which is not deliberately clear (Picheny et al., 1986). The duration and intensity of consonants tends to be enhanced in clear speech in comparison to the same measures for vowels (Fairbanks and Miron, 1957; Paul-Brown and Yeni-Komshian, 1988; Picheny et al., 1986), which again would seem to be an efficient manipulation, given that studies of misperception have shown that consonants are more often misperceived than vowels (Bond and

Garnes, 1980).

At the syntactic level, clear speech contains explicit cues to syntactic structure, such as complementisers and relative pronouns in positions where these are optional (Valian and Wales, 1976). Again, there is long-standing evidence from perceptual studies that listeners find it easier to process sentences which do contain such syntactic markers than sentences which omit them (Fodor and Garrett, 1967; Hakes and Cairns, 1970; Hakes and Foss, 1970; Hakes, 1972).

The present project has shown that clear speech also contains cues to the presence of a word boundary. As Cutler and Butterfield (1990) showed, speakers producing deliberately clear utterances tend to pause at word boundaries, and lengthen pre-boundary syllables. Once again, by marking word boundaries speakers use a tactic which will be of assistance to listeners, given that misperceptions of word boundary placement do occur both in spontaneous speech perception and in laboratory studies (Cutler and Butterfield, 1992). Indeed, Maassen (1986) has reported that low-intelligibility speech (produced by deaf speakers) can be rendered significantly more intelligible by insertion of pauses at word boundaries.

Moreover, speakers distinguish between different types of word boundaries: for instance, pausing tends to be significantly longer at boundaries which precede weak syllables than at boundaries which precede strong syllables. This too is likely to confer a perceptual advantage, given that English listeners apply a strategy of segmenting speech which assumes that strong syllables are very likely to be word-initial but weak syllables are not (Cutler and Norris, 1988; Cutler and Butterfield, 1992). The clear speech thus contains stronger cues to those boundaries which the normal perceptual strategies are most likely to overlook.

In agreement with this, the present study showed that in clear speech more changes are applied to weak syllables than to strong syllables. In part this fact is undoubtedly due to the vowel “reduction” which is such a consistent feature of weak syllables in English. Reduced vowels can always be unreduced, i.e. turned into full vowels; in this sense weak syllables offer greater oppor-

tunity for alteration from normal to clear speech. Furthermore, since weak syllables are far more likely than strong syllables to be omitted in misperceived speech (Bond and Garnes, 1980), it again makes sense for speakers to obviate this difficulty for listeners by, where possible, turning weak syllables into strong. But although such effects did occur in our subjects' productions, they were not sufficient to provide an explanation for all of the differences we observed between strong and weak syllables: in fact, the majority of weak syllables in the relevant experiments (1 to 3) contained reduced vowels in all three productions. Thus the present report further confirms the conclusion of Cutler and Butterfield (1990): in clear speech, more attention is paid to marking those word boundaries which are most likely to be overlooked. Most of the speaker's effort, however, is applied to pause insertion at boundaries and lengthening of pre-boundary syllables. The effects which we have reported here for post-boundary syllables, though consistent, are tiny in comparison with the durational effects described in our earlier report.

As we pointed out in the introduction, lexical segmentation is a major component of the listener's task; it is therefore no surprise that explicit word boundary cues contribute effectively to clear speech. Note that by providing explicit cues to segmentation, speakers are going beyond the mere application of global changes in speech rate or phonetic realisation; the alterations which they implement in the interests of added clarity are applied to the greatest extent where they are most needed. Indeed, our various measurements provide direct evidence for this observation. If differences between baseline and clear speech were due simply to global effects, all syllable types should show a similar pattern of alteration; but as our results show, this is not the case. Nor is it the case that the differences are due to global effects which happen to affect strong and weak syllables differently. If this were so, then the differences between strong and weak syllables would be similar irrespective of position relative to a word boundary; but again, they are not. Take, for instance, the durational data. After the boundary, weak syllables tend to be lengthened to a greater degree than strong; but before the boundary, the degree

of lengthening depends on what follows the boundary. In Experiments 1 and 2, when the post-boundary syllable was weak the pre-boundary syllable was usually strong, and vice versa; Cutler and Butterfield (1990) found that pre-boundary syllables were lengthened more before weak syllables, which in effect means that it was the *strong* pre-boundary syllables which showed more lengthening. Thus the strong-weak difference for pre-boundary syllables in Experiments 1 and 2 is in the opposite direction from the difference for post-boundary syllables – direct evidence against a global explanation of this difference.

The point can be further illustrated by computing correlation coefficients for the measures taken on pre- and post-boundary syllables. We have done this for the durational measures, firstly because these were the measures which (as reported by Cutler and Butterfield (1990)) showed the largest effects, and secondly because the between-syllable pause contributes a third durational measure. The coefficients of correlation between each pair of the three absolute measures of duration (across all subjects and items) are shown in Table 1.

Table 1
Correlations between durational measures (Pearson product-moment correlations, $df = 59$ for Experiments 1 and 2, 119 for Experiments 3 and 4)

	Pre-boundary syllable and pause	Pre-boundary syllable and post-boundary syllable	Pause and post-boundary syllable
Baseline			
Experiment 1	0.42, $p < 0.001$	-0.31, $p < 0.02$	N.S.
Experiment 2	N.S.	-0.30, $p < 0.02$	N.S.
Experiment 3	N.S.	N.S.	0.45, $p < 0.001$
Experiment 4	N.S.	N.S.	N.S.
First repetition			
Experiment 1	0.46, $p < 0.001$	N.S.	N.S.
Experiment 2	0.46, $p < 0.001$	-0.26, $p < 0.05$	N.S.
Experiment 3	0.64, $p < 0.001$	0.29, $p < 0.005$	0.21, $p < 0.03$
Experiment 4	0.67, $p < 0.001$	N.S.	N.S.
Second repetition			
Experiment 1	0.39, $p < 0.002$	N.S.	N.S.
Experiment 2	0.44, $p < 0.001$	N.S.	N.S.
Experiment 3	0.56, $p < 0.001$	0.31, $p < 0.005$	N.S.
Experiment 4	0.67, $p < 0.001$	N.S.	N.S.

Several predictions can be tested against the correlation pattern. Firstly, our claim is that pausing and pre-boundary lengthening are both word boundary cues in clear speech. Therefore, the duration of the pre-boundary syllable and the duration of the pause should be highly correlated in clear speech. The first column of Table 1 shows that this is true, without exception. However, the effects are specific to clear speech, as is shown by the fact that these measures are *not* correlated in the baseline utterances (with the exception of Experiment 1, in which subjects knew of the clear speech requirement when producing the baseline utterances). Our claim predicts, secondly, that there should be no direct correlation between duration of the post-boundary syllable and either of the other measures, since what is determining the pattern of results is, in our view, the phonological strength (essentially, the full or reduced vowel quality) of the post-boundary syllable, and this factor is not directly related to duration.

Other hypotheses, however, predict differently. If the lengthening effects noted by Cutler and Butterfield (1990) were simply due to global slowing of rate of speech in some sentences, then all durational measures should be correlated with one another in clear speech. Alternatively, one could argue that there might be a pressure in clear speech towards equalisation of syllable rate ("staccato" speech); this would predict *compensatory* effects, whereby pauses would be longer adjacent to shorter syllables and shorter adjacent to longer syllables – i.e., there should be a negative correlation between pause duration and either one or both of the syllable duration measures.

Clearly, the pattern of correlations does not support these latter predictions. It is not the case that all durational measures are correlated in clear speech: only the duration of pauses and pre-boundary syllables show a consistent relationship. Nor is there any sign of compensatory effects, i.e. negative correlations. Again, however, the prediction of our claim is supported: the duration of post-boundary syllables is not reliably correlated with either of the other measures. Twelve of the sixteen correlation coefficients involving post-boundary syllables in clear speech are insignificant. (Of the other four, two are relationships

present in the baseline and first repetition which disappear in the second repetition; the remaining two concern the same correlation, between the two syllable measures of Experiment 3. All four are much weaker than the predicted correlations in the first column.) We conclude, therefore, that the effects we have observed are definitely not global, but specific to word boundaries. In clear speech, speakers attempt to make the listener's task easier in a number of ways, and one of these is by marking word boundaries, especially boundaries which might otherwise be hard to detect.

Finally, note that in one respect our work may seem to have failed to replicate previous studies: Cooper and Paccia-Cooper (1980) reported that word frequency affected the degree to which speakers were willing to obscure word boundaries by the application of phonological assimilation rules, in that the rules were less likely to be applied when the word after the boundary had a low frequency of occurrence. We found no systematic effects of frequency, nor of the arguably related variables of predictability and word class. Cooper and Paccia-Cooper's studies did not explicitly vary the strong/weak contrast, so it could be hypothesised that they inadvertently picked up the same effect which we have found, by choosing low frequency words which happened to begin with weak syllables. However, this is not so: Cooper and Paccia-Cooper found that the assimilation rules were highly likely to be applied across boundaries preceding high-frequency function words, and these are precisely the words most likely to be realised as weak syllables. We suggest, instead, that Cooper and Paccia-Cooper's experiments tell us about effects which occur in normal speech; our experiments indicate that at least some of these effects will be nullified in clear speech.

In conclusion, we reiterate the point that speakers who are trying to be clear go about it in a very efficient way. This suggests that speakers base their clear speech strategies on a good model of the listening process. Otherwise, there is perhaps nothing very surprising about clear speech. Our work, in short, contributes to what is now a substantial body of evidence that speech which is intended to be clear is indeed just that.

Acknowledgment

This research was supported by IBM UK Scientific Centre. We are grateful to Brian Pickering for valuable discussions throughout the project, and to John Holdsworth and Brit van Ooyen for technical assistance. The *F0* results were previously reported to the Autumn Meeting of the Institute of Acoustics (UK), November 1990.

Appendix A. Experimental materials

For each experiment the six pairs of target sentences are listed, with the critical boundaries marked “/”, and the members of each pair labeled 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 awo 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

- Z.S. Bond and S. Garnes (1980), "Misperceptions of fluent speech", in *Perception and Production of Fluent Speech*, ed. by R. Cole (Erlbaum, Hillsdale, NJ), pp. 115-132.
- Z.S. Bond and L.H. Small (1983), "Voicing, vowel and stress mispronunciations in continuous speech", *Perception & Psychophysics*, Vol. 34, pp. 470-474.
- C.P. Browman (1978), "Tip of the tongue and slip of the ear: Implications for language processing", *UCLA Working Papers in Phonetics*, Vol. 42, pp. 1-149.
- S. Butterfield and A. Cutler (1990), "Intonational cues to word boundaries in clear speech?", *Proc. Institute of Acoustics*, Vol. 12, Part 10, pp. 87-94.
- 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 and J.M. Paccia-Cooper (1980), *Syntax and Speech* (Harvard Univ. Press, Cambridge, MA).
- A. Cutler and S. Butterfield (1990), "Durational cues to word boundaries in clear speech", *Speech Communication*, Vol. 9, Nos. 5/6, pp. 485-495.
- A. Cutler and S. Butterfield (1992), "Rhythmic cues to speech segmentation: Evidence from juncture misperception", *J. Memory Language*, Vol. 31.
- A. Cutler and D. Carter (1987), "The predominance of strong initial syllables in the English vocabulary", *Comp. Speech Language*, Vol. 2, pp. 133-142.
- A. Cutler and C.E. Clifton (1984), "The use of prosodic information in word recognition", in *Attention and Perform-*

- mance X: Control of Language Processes, ed. by H. Bouma and D.G. Bouwhuis (Erlbaum, Hillsdale, NJ), pp. 183–196.
- A. Cutler and D.G. Norris (1988), "The role of strong syllables in segmentation for lexical access", *J. Experimental Psychology: Human Perception Performance*, Vol. 14, pp. 113–121.
- G. Fairbanks and M.S. Miron (1957), "Effects of vocal effort upon the consonant-vowel ratio within the syllable", *J. Acoust. Soc. Amer.*, Vol. 29, pp. 621–626.
- J.A. Fodor and M.F. Garrett, (1967), "Some syntactic determinants of sentential complexity", *Perception Psychophysics*, Vol. 2, pp. 289–296.
- W.N. Francis and H. Kučera (1982). *Frequency Analysis of English Usage* (Houghton Mifflin, Boston, MA).
- 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.
- D.T. Hakes (1972), "Effects of reducing complement constructions on sentence comprehension", *J. Verbal Learning Verbal Behavior*, Vol. 11, pp. 278–286.
- D.T. Hakes and H.M. Cairns (1970), "Sentence comprehension and relative pronouns", *Perception Psychophysics*, Vol. 8, pp. 5–8.
- D.T. Hakes and D.J. Foss (1970), "Decision processes during sentence comprehension: Effects of surface structure re-considered", *Perception Psychophysics*, Vol. 8, pp. 413–416.
- D.H. Klatt (1975), "Voice onset time, frication and aspiration in word-initial consonant clusters", *J. Speech Hearing Res.*, Vol. 18, pp. 686–706.
- A. Malécot (1958), "The role of releases in the identification of released final stops", *Language*, Vol. 34, pp. 370–380.
- B. Maassen (1986), "Marking word boundaries to improve the intelligibility of the speech of the deaf", *J. Speech Hearing Res.*, Vol. 29, pp. 227–230.
- D. Paul-Brown and G. Yeni-Komshian (1988), "Temporal changes in word revisions by children and adults", *J. Speech Hearing Res.*, Vol. 31, pp. 630–639.
- M.A. Picheny, N.I. Durlach and L.D. Braida (1985) "Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech", *J. Speech Hearing Res.*, Vol. 28, pp. 96–103.
- 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. Speech Hearing Res.*, Vol. 29, pp. 434–446.
- K. Shafer-Vincent (1983), "Pitch period detection and chaining: Method and evaluation", *Phonetica*, Vol. 40, pp. 177–202.
- 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.