The effect of delaying auditory feedback of selected components of the speech signal

PETER HOWELL
University College London, London, England

Can the disturbance that occurs in delayed auditory feedback tasks be explained by problems caused by the amplitude contour of the displaced signal, or is information about the position of the articulators used for feedback control? To test these alternative hypotheses of the effects of delayed auditory feedback, the effects of varying the amplitude contour of the delayed signal were assessed when phonetic information in the delayed signal was degraded. Two experiments are reported in which the speech signal is split into two bands and each band is delayed by different times. The amplitude of one or both of these bands was varied in Experiment 2. Performance was not as disrupted when only one band was delayed as when both were. In Experiment 2, it was found that increasing the amplitude of the band that was not delayed did not disrupt performance, whereas increasing the amplitude of the delayed band did disrupt performance, particularly at a delay of 200 msec. The results are interpreted as showing that the amplitude contour of the displaced signal causes control problems with speech in delayed auditory feedback tasks.

There is no doubt that altering the auditory feedback of a speaker's voice causes speech control to suffer. There is, however, no consensus about how these effects should be explained. On the one hand, they have been regarded as supporting a process of feedback control during speech production (Fairbanks, 1955; Lee, 1950) On the other, at least some of these effects can be accounted for more satisfactorily in other ways (Borden, 1979; Howell, Powell, & Khan, 1983; Lane & Tranel, 1971). This paper contributes to this issue by examining whether the effects of changing auditory feedback by delaying it support a feedback-control account of speech production or not. First, the effects of delayed auditory feedback (DAF) are described.

The most apparent effect of DAF is that it causes normal speakers to drawl and stutter and produce all types of speech errors (Lee, 1950). Speakers also increase voice pitch and the intensity of their speech (Fairbanks, 1955). These disruptive effects vary with the amount by which the speech signal is delayed such that speech errors and slowing are at a maximum when the speech is delayed by about 200 msec, and voice pitch and intensity asymptote at about this same delay (Fairbanks, 1955).

These findings were originally interpreted as support for a feedback-control account of normal speech production (Fairbanks, 1955; Lee, 1950). Thus, when the feedback is deranged, the speaker's normal means

of controlling his voice are deranged too. The peak on errors and time at a 200-msec delay was thought to arise because it was supposed that articulation was planned for segments about 200 msec in length. Hence, when speech is delayed by 200 msec, feedback is not received until the speaker has started producing the next 200-msec-long unit. The subject thinks that the feedback he is hearing is from the unit he is currently producing but, in fact, it is from the previous one. Thus, the effects of altered feedback and, in particular, maximum disturbance at a delay of about 200 msec are consistent with an account that holds that speech is under feedback control. It follows from this account that the programming unit in speech production is of the same order of length as the delay that produces maximum disruption, and, that since syllables are about 200 msec long, they might be the units used in the control of speech production.

Later theorists have considered that the disruption observed under altered feedback may not provide evidence for a feedback-control account of speech production. Instead, alterations to feedback may force the speaker to pay attention to the changed feedback (Borden, 1979; Lane & Tranel, 1971). Borden (1979) has argued that the speaker tries to change his speech so as to correct for the alterations that have been made to it. On Borden's explanation, if feedback is delayed, the speaker drawls in an attempt to bring the speech he hears back into synchrony with what he is producing. Borden's account can also explain effects found when feedback is altered in ways other than delay. The findings that if a speaker's speech is amplified he reduces his speech level (Siegel & Pick, 1974) and that he shifts the fun-

damental frequency of his voice in the direction opposite to that to which his vocal output is frequency-shifted (Elman, 1981) are consistent with this account. If Borden is correct, studies of altered feedback would not have many implications for the process of normal speech control, because the speaker would not normally experience such alterations. Lane and Tranel (1971) explain the changes in speech when the level of auditory feedback is altered as being brought about by changes the speaker makes when trying to maintain the intelligibility of his speech. They argue that such changes are effected for the sake of listeners, not to maintain effective feedback for the speaker himself.

If these latter accounts are correct, experimental studies of altered auditory feedback would not provide evidence that feedback is used in the moment-to-moment control of speech production. Can, then, the effects of altered feedback be ignored because they are experimental curiosities? Obviously not, because altered feedback techniques have found practical applications. Applications of changes in feedback level include the diagnosis of auditory dysfunction (Lane & Tranel, 1971), and, of the effects of DAF, the detection of auditory malingering (Yates, 1963) and the treatment of stuttering (Leith & Chmiel, 1980).

Hence, it is still essential to know why the effects of altered feedback occur, not least for the role such explanations might play in explaining auditory and speech dysfunction. Also, if the effects of altered feedback do not depend on the speaker's continuing to use the recurrent signal for feedback control, the explanation of normal speech control by feedback regulation is brought into question, too, since altered feedback studies are an important line of support for such explanations. Explanations of the effects of altered feedback that deny that the speaker continues to use the altered signal for feedback control have been offered by Lane and Tranel (1971), for the effects of level, and by Howell et al. (1983), for the effects of DAF. Lane and Tranel (1971) explained the effects of level changes as an attempt on the part of the speaker to maintain a constant signal-to-noise ratio. Howell et al. (1983) have argued that the effects of DAF are just one example of difficulties experienced in performing a serial motor task when some distracting event is going on at the same time.

The tenet of the latter hypothesis is that disruption to any motor act will occur when some other event is taking place concurrently (but not in synchrony) with it. The disruption that is observed in DAF tasks arises because the delayed signal is going on concurrently with the speech rather than being attributable to the speaker's continuing to try to use the recurrent signal for feedback control. What is it about hearing a delayed version of speech that is so disruptive for speech control, and why should this be so marked when the two events occur at an offset of about 200 msec? Howell et al. (1983) proposed that the delayed speech signal was so effective at disrupting performance because of the relationship between the amplitude contour of the speech (which he continues to hear through bone conduction and because the attenuation afforded by the headphones is incomplete) and that of the delayed signal. There is a maximum in the amplitude contour associated with each syllable and a minimum at the end of each syllable (Fant, 1959). Thus, when speech is delayed by 200 msec, the amplitude of the delayed signal is beginning to rise while that of the currently spoken syllable is falling to a minimum. A signal delayed by 200 msec would be the most intrusive because it starts to rise as the syllable as vocalized is finishing. This would then cause most disruption to performance. Note that, because effects at the level of the syllable explain why so much disruption occurs at a delay of 200 msec, it does not follow that the syllable is the programming unit of speech production.

One implication of the hypothesis is that the phonetic identity of the speech is not necessary to produce the disruption observed in DAF tasks. Experimental evidence on this question has been provided by Howell and Powell (in press). In this experiment, the amplitude envelope of speech was calculated and this envelope alone delayed. The delayed-amplitude envelope was then used to gate a sawtooth wave on and off. What the speaker hears is the nonspeech sound delayed and having the same amplitude contour that the speech sound would have in standard DAF tasks. That is, the "feedback" has the same amplitude contour as speech in DAF tasks but none of the phonetic characteristics of the original speech. Howell and Powell (in press) reported that if a speaker is timed to speak a list when hearing this sort of feedback, a peak in the disturbance function occurs at a delay of 200 msec, as in DAF tasks. On the basis of this result, it appears that the spectral composition of speech is not essential to produce the disturbance noted in DAF tasks, but that the relationship between the amplitude contour of speech and that of the delayed signal is.

If the peak at a delay of about 200 msec occurs because of the amplitude contour of the displaced speech, it should be affected by alterations to the amplitude contour of the speech. It is this aspect of the hypothesis that the experiments to be reported were designed to examine. There are many ways that the amplitude contour of speech could be altered to test this prediction. The technique chosen in the present experiments was to filter the speech into two bands and to delay one of the bands relative to the other. This makes the amplitude peaks of the altered feedback less prominent than in "normal" DAF. In normal DAF, the direct signal is attenuated by the cushions of the headphones. The amplitude contour of

the delayed signal dominates, and this causes disruption to speech activity. By separating the speech into two bands and routing one band back directly and the other after a delay, the amplitude contour in each part overlaps and neither dominates. A disadvantage is that although, ideally, the amplitude contour of the feedback should be flat so as to be equally intrusive at all delays, it is not so when feedback is altered in this way and, also, the contour varies somewhat with delay. There are, on the other hand, three principal advantages in using this technique. First, filtering the speech into two bands and delaying one relative to the other makes the speech less intelligible and unlikely to be used for feedback control.[1] Taken in conjunction with the results of Howell and Powell (in press), any effect such feedback has, then, is unlikely to arise because the speaker listens to the feedback for its phonetic content. It is more likely that the disturbance observed with such a signal is attributable to the amplitude contour. Second, the amplitude contour of one band or the other can be altered independently by bringing the amplitude of one part up or reducing the amplitude of the other. It should be possible by appropriate manipulations of one or the other amplitude contour to cause the peak in the disturbance function to appear and disappear. This would not be possible if other techniques, such as volume limiting of the amplitude contour, were employed. Third, the bands have a specifiable time of occurrence relative to the speaker's original utterance.

In the first experiment, the intention was to see whether altering the amplitude contour by this technique removed the peak in the disturbance function [i.e., the function relating time to read a list (PT) to delay with a peak at a delay of about 200 msec]. The task was to read a list of vowels as fast as possible. In the experimental conditions, the speech was filtered into a low-frequency band (less than 550 Hz) and a band between 550 Hz and 3.5 kHz by electronic filters. One of these bands was fed back directly and one after a specified delay. The cutoff frequencies of the filters were chosen so as to separate out regions of energy concentration in the vowels chosen. The regions in which speech energy is concentrated are called formants, and by convention they are numbered from the lowest frequency (first formant, F1) up (F2, F3, etc.). Vowels were employed for two reasons. First, their formants move less over time than do the formants of consonants, and fixed filter cutoffs can be used. The low-pass filter passes F1 and the bandpass F2 and F3.[2] Second, vowels are appropriate for testing a second hypothesis. It is possible to argue that although the phonetic identity of speech is reduced by filtering it, the speaker is still able to extract information from each band to control performance. If this were so, phonetic information would still occur in the altered signal and the manipulation would not leave a signal varying essentially in its amplitude contour. Of course, this would not

explain the results of Howell and Powell (in press) referred to above. Nevertheless, this hypothesis is still worth testing, and this was done as follows. The vowels of English vary in tongue height and front/back position of the tongue in the vocal tract. Tongue height is inversely related to F1 frequency and front/back position to (F2-F1). If speakers are able to extract information from each band, they should be affected differently, depending on which part of the signal is delayed for different pairs of vowels. For example, such vowel pairs as /u/ and /a/ vary mainly in tongue height (i.e., vary mainly in F1 frequency), and they might be more disrupted when the low band (contains F1) is delayed but not when the high band is delayed. Vowel pairs such as /u/ and /i/ vary in front/back position but not in tongue height, so they differ in the frequency of F2. /i/ and /a/ vary in both F1 and F2 frequency. In addition, these vowel pairs differ with respect to the frequency of F3. For the vowels chosen, F3 frequency is directly related to F2 frequency—that is, either F2 and F3 both vary across the vowel pair or neither differs to any extent. In sum, the intentions of Experiment 1 are to see whether removing the amplitude modulations of speech in the way described reduces the peak in the disturbance function relative to normal DAF and to see whether speakers can use information in separate bands to control performance.

EXPERIMENT 1

Method

Subjects. Seven male subjects with no history of hearing loss were paid for participating in the experiment. They were all right-handed native English speakers.

Stimulus materials. Vowel sets were chosen which were differentiated by information in one of the two frequency bands. Three vowel sets were employed: /i/ and /a/, /i/ and /u/, and /a/ and /u/.

The vowels in the /i-u/ set have the same linguistic specification for tongue height (both are high) but have different front/back positions in the vocal tract (/i/, front; /u/, back). This set should have similar F1s but different F2s (F2 of /u/ is lower than that of /i/) and corresponding differences in F3 from those in F2. The vowels /u/ and /a/ have different heights (/u/, high; /a/, low), but both are back vowels. Thus, this set should differ in F1 but not in F2/3. The vowels /i/ and /a/ have different tongue heights and different front/back positions of the tongue. Consequently, they have different F1s and (F2-F1)s as well as corresponding differences in F3 to those in F2.

Feedback conditions. The speakers were to speak these lists under three feedback conditions: F1 delayed, F2/3 not; F2/3 delayed, F1 not; "normal" DAF. Initial transduction of the speech signal and playback of the signal to the speaker/hearer was always performed on the same equipment. The speech was picked up by a moving coil microphone (Beyerdynamic M 88 N) and amplified by a Revox tape recorder. After the signal was filtered and delayed (see below), the signal was fed through a second Revox, operating as an amplifier, and presented to the speaker over HD414X headphones. The frequency response of the headphones measured on real ears was flat to ±7 dB from 250 Hz to 10 kHz. Spectrographic analysis of the speakers' normal utterances showed that F1s of all the vowels of the speakers used in this paper could be separated from the F2s by low-pass filtering at 550 Hz (48 dB/octave). The high-frequency feedback band consisted of the speech

bandpass filtered between 550 Hz and 3.5 kHz (48 dB/octave). If the speaker was performing the F1 delay condition, the low-pass filter output was sent to a PDP-11/10 computer that was programmed to act as a variable delay line. This worked by digitizing the speech at a 10-kHz rate and starting the output (also at 10 kHz) after the specified delay. All computer outputs were low-pass filtered at 3.5 kHz. If F1 was being delayed, the high-frequency bandpass filter output was led directly to the second tape recorder, where the signals were mixed. If the high-frequency bandpass filter output was to be delayed, it was routed to the computer, and the low-pass filter output was led directly to the tape recorder that was mixing and amplifying the signal before it was presented to the speaker. When all of the speech was to be delayed, the speech was filtered in the manner outlined above but both low-pass- and bandpass-filtered outputs were combined before they were led to the computer for delaying. The filtering was performed even though all the signal was to be delayed in order to make the conditions comparable.

Level. The level of feedback is an important determinant of the DAF effect (Butler & Galloway, 1957). To determine the level the feedback was fed back at, a calibrated pink noise source at 58 dB (Ivie IE 20 B) was used to measure the output level of the apparatus. The noise source was substituted for the vocal input, and the level was measured at the headphone output. Under normal DAF, the signal was 26 dB higher than at input. The low-pass-filtered signal alone was 22 dB up and the bandpass-filtered output was 24 dB up. These and all other levels in this paper were measured on a Brüel and Kjaer meter with an A weighting network. The signals were left at this level throughout the experiment.

Procedure. Each subject received all seven delays (0, 0.1, 0.15, 0.2, 0.25, 0.3, and 0.5 sec) for each of the vowel sets under each of the feedback conditions. The speaker performed all of the delays under one of the feedback conditions before he performed any delay in the other feedback condition. Six subjects received different orders of the three feedback conditions and the seventh repeated one of the orders. The order of delays within each feedback condition was determined by a Latin square. A different Latin square was used in each of the feedback conditions. Once the feedback condition and delay had been selected, the subject received all three vowel sets. Each subject performed the vowel sets in the same order throughout the experiment, but the order of the vowel sets was counterbalanced across subjects.

Because speakers might be able to repeat some sequences faster than others, randomized sequences for each of the vowel pairs were generated in the following way: There are eight possible three-tuples for a binary event. A sequence of 162 items was computed which contained an equal number of each of these three-tuples (each occurred 20 times). There were equal numbers of two-tuples (i.e., alternations and repetitions) and an equal number of each of the vowels.

The subjects were timed with a stopwatch to repeat 40 items from the list of 162. All parts of the list were read equally often across speakers. The subject was told to speak the list as fast as possible but to make no errors. The experimenter monitored the lists to make sure they were spoken correctly. (This same experimenter also did the timing in Experiment 2.) If speakers misarticulated any item, the list was repeated. Errors were phonetic errors, and not quality differences. Vowel lengthening or pausing were not counted as errors. Lists had to be repeated less than 5% of the time because of errors. When a list had to be repeated, this was done at once. Lists were terminated immediately after an error was detected to prevent subjects from practicing the whole list.

Results and Discussion

The mean list performance time (PT) across subjects against delay are presented separately for each feedback condition in Figure 1.

It is clear from Figure 1 that delaying all the signal is much more disruptive than delaying F1 or F2/3 alone. List PT is less for the two part-delay condi-
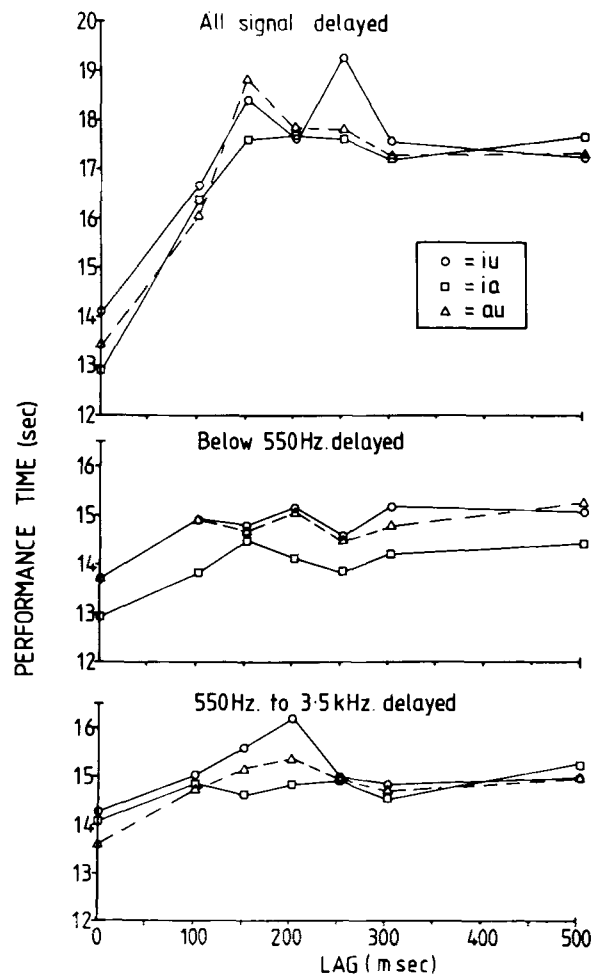

Figure 1. List reading time versus delay for each of the three feedback conditions in Experiment 1. The feedback condition is given by the label at the head of each graph. The symbol of the points refers to the three vowel sets, and the vowel set can be identified from the box in the inset.

tions. This impression was confirmed by analysis of variance. The factors in the analysis were subjects (7), vowel set (3), feedback delay condition (3), and delay (7 levels). The main effect of feedback delay condition was significant [$F(2,12) = 11.0$, $p < .01$], and a post hoc comparison showed that the part delay conditions had shorter PTs than the condition in which all the signal was delayed [$F(1,12) = 9.25$, $p < .05$].

For the all-delay condition, there is most slowing at around 200 msec, as has been reported by other workers (e.g., Butler & Galloway, 1957; Fairbanks, 1955). In the part-delay conditions, there is some increase in PT, but the peak is much less clear. These impressions also receive statistical support. Analysis of variance of the all-delay conditions with the factors of vowel set and delay resulted in a significant effect of delay, which orthogonal breakdown showed was because the linear [$F(1,6) = 11.7$, $p < .025$] and

quadratic components [F(1,6) = 6.75, p < .05] were significant. The quadratic trend indicates that the function is peaked, but the peak is not necessarily at delays comparable to those expected from DAF experiments (i.e., round about 200 msec). A polynomial regression of mean PT over delays was performed, and the predicted quadratic function was inspected to see where the peak lay. The delay at which maximum disruption occurred was 310 msec. The analysis of variance of the part-delay conditions with the same factors revealed that only the linear trend of the delay effect was significant [F(1,6) = 6.78, p < .05]. There were no interactions with vowel set, so, statistically speaking, the subjects were performing similarly with the different vowel sets. Thus, since the quadratic trend was not significant, in these conditions there was no significant peak.

It was hypothesized in the introduction that if speakers were able to use parts of the speech signal to control speech, speech control should be affected by whether the informative band was the one delayed or not. However, there is no indication that delaying the informative part of the signal is more disruptive than delaying the noninformative part of the signal. This would have been so if the interaction between vowel condition and feedback condition had been significant. The only effect of vowel set was its main effect, with the /i-u/ set being shorter than the /u-a/ set, which, in turn, was shorter than the /i-a/ set [F(2,12) = 7.0, p < .01].

To sum up the findings of Experiment 1, the disturbance function is not peaked when part of a signal is delayed, as compared with when all the signal is delayed. The lack of a peak in the part-delay conditions cannot be attributed to the speaker's using information in the direct band to control his speech. It appears, then, that the peak is removed because the technique alters the amplitude contour of the feedback in the part-delay condition, as predicted by the hypothesis outlined in the introduction (Howell et al., 1983). In the next experiment, the hypothesis is tested further. The peak in the disturbance function is attributed to amplitude contours displaced by about 200 msec relative to the speech being most disruptive on performance. If, in the part-delay conditions, the level of the delayed part is increased, that event should be more dominant and a peak should appear in that condition too, even though the phonetic content is degraded. This prediction was tested in the next experiment.

EXPERIMENT 2

Method

Subjects. Two groups of subjects, each consisting of eight males, were employed. They were all right-handed, native English speakers, and they were paid for their services.

Stimulus materials and sequences. The /i-u/ vowel set and stimulus sequence of Experiment 1 were employed. The other vowel pairs used in Experiment 1 were not included because there were no important differences between the vowel sets.

Feedback conditions. One group of eight subjects performed the all-delay condition and the F1 delayed conditions of Experiment 1, and the other group of eight subjects performed the all-delay and F2/3 delayed conditions of Experiment 1. The feedback conditions were the same as in the previous experiment, with the addition that the level of feedback was varied in the manner described next.

Level. In all feedback conditions, the speaker/hearer received F1 at three levels and F2/3 at three levels. There was a factorial combination of F1 and F2/3 levels, giving nine conditions in total. F1 was either switched off at the Revox, set to amplify a 58-dB calibrated pink noise source by 10 dB (Ivie IE 20 B) or to amplify it by 20 dB. The level of the voice when the Revox channel was off was somewhat less than baseline (0 dB gain) because of the attenuation offered by the ear cushions. Calibrated settings (off, +10 dB, +20 dB) for the F2/3 channel were obtained in a similar manner.

Procedure. Subjects were assigned to the group who had F1 delayed in the part-delay condition (Group 2a) or to the group who had F2/3 delayed (Group 2b). The two groups were treated the same, except in respect to which part of the signal was delayed in the part-delay condition.

Only four delays were employed in this experiment (0, 0.1, 0.2, and 0.3 sec). The subject received all conditions at one delay before he received those at another delay. Within each of the delay groups, a subgroup of four subjects had their delay orders determined by one Latin square and the other subgroup of four subjects by a different Latin square. Once a delay time had been determined for a particular subject, he either received all his all-delay conditions and then his part-delay conditions or vice versa. The order all/part was counterbalanced across subjects. Once the feedback condition had been selected, the subject received all the F2/3 levels at a particular F1 level before he received the F2/3 levels at the next F1 level, and so on. The order of F1 and F2/3 levels was counterbalanced across subjects. Other details of the experiment were as for Experiment 1.

Results and Discussion

The mean results across subjects are presented in Figures 2 to 5. In Figure 2, the results of Group 2a (F1 part-delayed) are presented with PT against delay. The results of the all-delay condition are on the left of the figure and those of the part-delay condition, on the right. Conditions in which the level parameters are equivalent are presented in corresponding rows. At the top of each pair of graphs is presented the F1 level, and the parameter of the connected points is the F2/3 level given in the caption in the box. Similar plots are presented for Group 2b in Figure 3 (F2/3 part-delayed). The difference between the plots in Figures 2 and 3 is that the F2/3 level is the caption of the row in Figure 3, and the points connected together and given the same symbol are combined by level of F1. The data of Group 2a are replotted in Figure 4 in the same way as those of Group 2b and the data of Group 2b are replotted in Figure 5 in the same way as those of Group 2a.

As an indication of the reliability of the data, some of the conditions in the all-delay feedback condition can be compared with conditions that should be
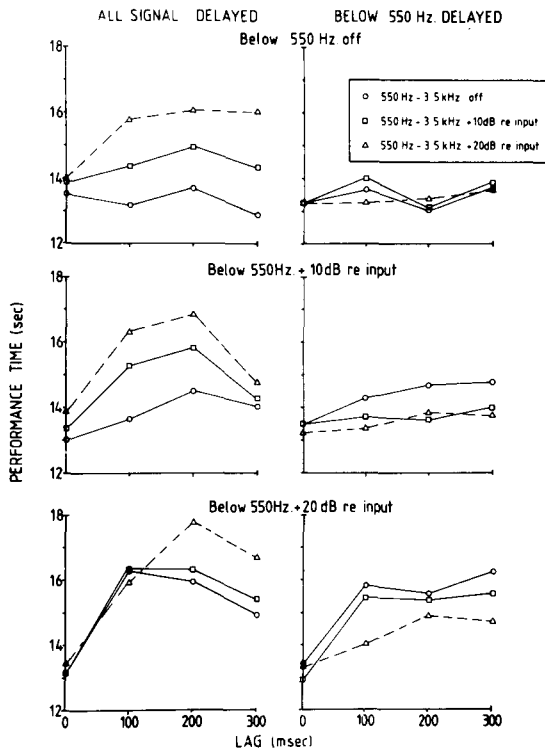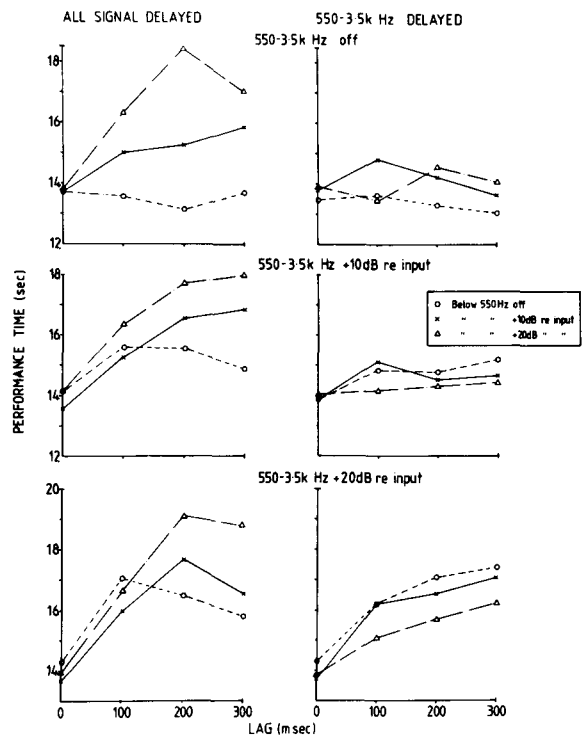
Figure 2. List reading time versus delay for Group 2a of Experiment 2. The level of F1 (below 550 Hz) is given at the head of each pair of graphs in a row. The level of F2 is indicated by the symbol, and this can be identified from the box on the right. The graphs in the left column come from the all-delay feedback conditions and those on the right from the part-delay feedback conditions.

equivalent in the part-delay feedback condition. When the feedback is not delayed at all, or when the delayed part in the part-feedback condition is switched off, the speaking situation is the same as it is in the corresponding all-delay condition, in which the band is delayed but switched off. In Figures 2 and 3, the far left points of the graphs in each row should correspond, which they do. When F2/3 is switched off in Figure 2 or when F1 is switched off in Figure 3, the PTs are comparable as well.

All-delay condition. Inspection of the all-delay conditions of Group 2a (Figures 2 and 5) and Group 2b (Figures 3 and 4) shows that PT increases whether the F1 or the F2/3 level is increased for both groups. Butler and Galloway (1957) reported that the disturbance functions become more "peaky" as level of the delayed signal is increased. Statistics on the all-delay conditions support this. An analysis of variance with factors of groups (2), subjects (8 per group), F1 level (3), F2 level (3), and delay (5) was performed. There was no difference between groups except for an interaction between delay, F1 level, and groups [F(6,84) = 2.8, p < .05]. There is no obvious explanation of this. The absence of any substantial differences between groups indicates that the subjects in

the two groups were performing similarly, although, because of the interaction, not identically. The effects of F1 level [F(2,28) = 20.6, p < .001] and of F2/3 level [F(2,28) = 25.2, p < .001] on PT indicate the increase in PT with increased level. The main effect of delay demonstrates that PT differs across delays [F(3,42) = 20.6, p < .001]. Both the linear and quadratic components of the delay effect were significant [F(1,14) = 17.9, p < .001, and F(1,14) = 40.8, p < .001, respectively]. Other factors interacted with delay, which indicates that only certain of the disturbance functions are peaked (a partial ANOVA table is given in Table 1).

The F1 level interacted with the quadratic component of the delay effect [F(1,14) = 17.8, p < .001] and the F2/3 level interacted with the quadratic component of the delay effect, too [F(1,14) = 13.1, p < .005]. These effects suggest that increasing the level of either component increases the peakiness of the function when it is delayed. A polynomial regression was used to determine whether a significant peak had occurred for each of the F1 and F2 levels for each of the groups. To see whether a significant peak had occurred, a goodness-of-fit test was performed to ascertain whether a trend of higher order than linear

Figure 3. List reading time versus delay for Group 2b of Experiment 2. The level of F2 (550 Hz-3.5 kHz) is given at the head of each pair of graphs in a row. The level of F1 is indicated by the symbol, and this can be identified from the box on the right. The graphs in the left column come from the all-delay feedback conditions and those on the right from the part-delay feedback conditions.
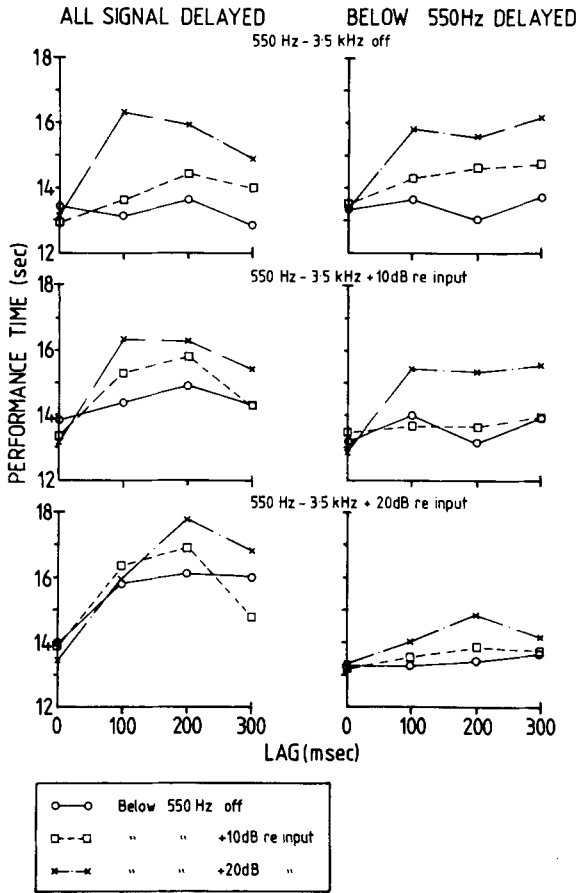
Figure 4. All-delay feedback conditions for Group 2a replotted in the way the all-delay feedback conditions of Group 2b were in Figure 3.

should be taken into account in fitting the function. An F statistic was computed in which the numerator was the sum of squares of all polynomials of higher order than linear and the denominator was the residual sum of squares of the fit to the highest degree polynomial. This statistic showed that there tended to be a significant peak when F1 and/or F2/3 was at a high level. This extends Butler and Galloway's (1957) observation by showing that the level of either of these components increases the peakiness if it is delayed. The delay at which the peak occurred for the conditions in which the peak was significant are presented for each group in Table 2. The peak occurred at delays ranging from 180 to 260 msec, which is roughly comparable to the delays reported in other studies (e.g., Butler & Galloway, 1957; Fairbanks, 1955).

Part-delay conditions. The findings in the part-delay conditions for Groups 2a and 2b can be summarized by saying that increasing the level of the direct (real-time) component reduces PT (Figures 2 and 3) and increasing the level of the delayed component increases PT (Figures 4 and 5). The PTs are

not as long, and nor are the disturbance functions as "peaky," as in the all-delay condition.

To confirm these observations statistically, the data from the two part-delay conditions of the two groups were first analyzed separately, and no differences between the groups were noted. The analysis of variance for the combined groups is reported. The factors in the analysis of variance were groups (2), subjects (8 per group), level of the delayed component (3), level of the nondelayed component (3), and delay (4). There were no significant effects of groups at all. The main effects of delayed level [$F(2,28) = 130.9$, $p < .001$], of non-delayed level [$F(2,28) = 6.3$, $p < .01$], and of delay [$F(3,42) = 7.4$, $p < .001$] were significant, as were all the two-way interactions between these factors (Table 3). Inspection of the means for the factors that had a significant main effect shows an increase in PT with level of the delayed component and delay but a decrease in PT with the level of the nondelayed component. The linear [$F(1,14) = 9.5$, $p < .01$] and quadratic [$F(1,14) = 5.3$, $p < .05$] components of the delay effect were significant.

In addition to the main effect of the quadratic component of delay, the quadratic component of the delay effect interacted with level of the delayed component, too [$F(1,14) = 17.3$, $p < .01$]. This effect
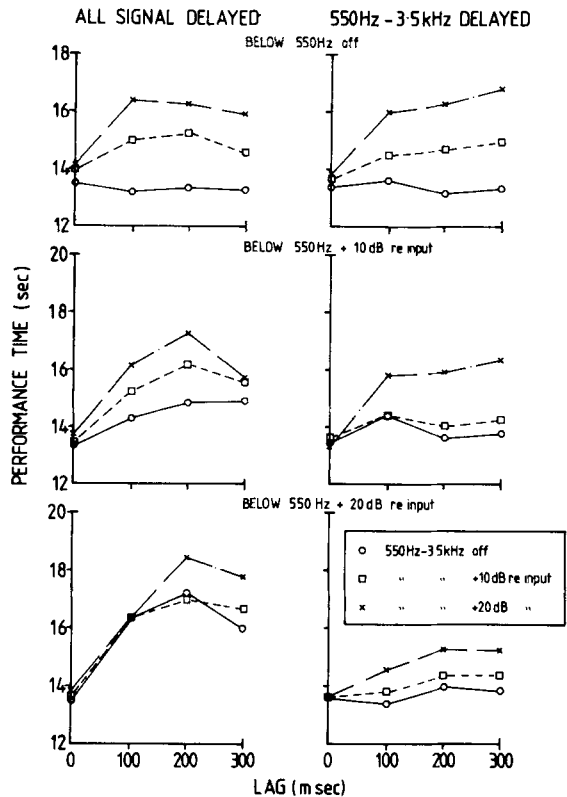

Figure 5. All-delay feedback conditions of Group 2b replotted in the way the all-delay feedback conditions of Group 2a were in Figure 2.

Table 1
Analysis of Variance Table (Partial) for the All-Delay Conditions of Experiment 2

Source	SS	df	MS	F	p <
F1 Level	218.641	2	109.32	20.6	.001
Error	148.5	28	5.304		
F2 Level	208.875	2	104.438	25.24	.001
Error	115.844	28	4.137		
Lag	513.922	3	171.307	18.8	.001
Error	382.719	42	9.112		
F1 Level x F2 Level	38.734	4	9.684	5.07	.01
Error	106.906	56	1.909		
F1 Level x Lag	118.25	6	19.708	10.37	.001
Error	159.625	84	1.900		
F2 Level x Lag	42.672	6	7.112	3.41	.01
Error	175.016	84	2.084		

shows that the peak (quadratic trend) emerges as the level of the delayed component is increased. The quadratic trend did not interact with the level of the direct component. This suggests that the peak occurs when the level of the delayed component was increased but not when the level of the nondelayed component was increased. To ascertain whether this was so, a regression analysis was performed as with the all-delay condition. A peak was estimated for those conditions in which the quadratic component was needed to obtain a good fit, and these are presented in Table 4. The peaks range between 140 and 280 msec— that is, comparable to those in ordinary DAF tasks. A peak always occurred when the level of the delayed component was at its highest (+20 dB), but only once when it was at its lowest level. There was no noticeable tendency of the functions to peak as the level of the nondelayed component was increased.

GENERAL DISCUSSION

In Experiment 1, a peak in the disturbance function was obtained in normal DAF, but none was ob-

Table 2
Estimates of Delay (in Milliseconds) at Which Peak
Occurred for the Various F1 and F2/3 Levels in
the All-Delay Conditions of Experiment 2

		F1 Level		
		0	+10	+20
		Group 2a		
	0	n.s.	n.s.	185
F2/3 Level	+10	n.s.	180	190
	+20	220	180	220
		Group 2b		
	0	n.s.	n.s.	215
F2/3 Level	+10	180	260	260
	+20	185	215	260

Note—The delay at which the peak occurred was only estimated in conditions where regression analysis indicated that a quadratic component was needed to provide a good fit.

served when only part of the original speech signal was delayed. In Experiment 2, the effect of varying the amplitude of each of two frequency bands when both (normal DAF) or only one (part delay) of the bands was delayed was examined. A peak in the disturbance function occurred in this experiment when the amplitude of the delayed band (part) or bands (normal DAF) was at a high level. These findings are consistent with an account that holds that the effects of delayed feedback are associated with the amplitude contour of the delayed speech rather than with more traditional accounts, which maintain that the late arrival of phonetic information at control causes speaking problems. The other finding that specifically supports the hypothesis favored here is that nonspeech gated on by the amplitude contour of speech gives similar disturbance to that observed with DAF (Howell & Powell, in press).

What other effects of DAF are there that a satisfactory hypothesis should explain? There are numerous examples of disturbance, similar to that noted with delayed feedback, with activities other than speech (see Howell et al., 1983, for a review). Such findings are easily explained by the present hypothesis provided that a similar relationship between the activity's offset and the onset of feedback can be identified. The hypothesis predicts that the "feedback" in speech and nonspeech tasks does not have to arise from activities specifically associated with the activity itself. There are several informal examples that indicate that this is probably so. Black's (1951) interest in the DAF phenomenon with speech, for example, seems to have been promoted by his observation of problems recording engineers have in monitoring their speech when other people are speaking at the same time. Many musical tasks that are difficult to perform, such as hand-bell ringing, may be so because the performer has to synchronize his activity with respect to the offset of a sound produced by another performer. In both these examples, the intruding event would not be used in order to regu-

Table 3
Analysis of Variance Table (Partial) for the Part-Delay Conditions of Experiment 2

Source	SS	df	MS	F	p <
Nondelayed Level	16.711	2	8.355	6.3018	.01
Error	37.125	28	1.326		
Delayed Level	261.828	2	130.914	28.6602	.001
Error	127.898	28	4.568		
Lag	121.664	3	40.555	7.4425	.001
Error	228.859	42	5.449		
Nondelayed Level x Delayed Level	35.000	4	8.750	10.5554	.001
Error	46.422	56	0.829		
Nondelayed Level x Lag	18.227	6	3.038	3.279	.01
Error	77.820	84	0.926		
Delayed Level x Lag	87.461	6	14.557	9.1962	.001
Error	133.148	84	1.585		

late muscular control involved in the activity, since the "feedback" is not a consequence of the activity itself. The disruption probably arises because the event is particularly intrusive, as is the delayed signal in DAF tasks.

Another effect of DAF that any satisfactory theory must be able to account for is the difference in susceptibility of different subjects to the disruptive effects of DAF. Explanations of different susceptibility in traditional accounts have sought to identify differences between subjects with respect to reliance on feedback from different sources (e.g., tactile, proprioceptive, or auditory). There is little compelling evidence for such explanations (again, see Howell et al., 1983, for a review). Howell et al. predicted that the physical properties of the speaker's voice could account for differences in susceptibility to the effects of DAF. It is known that speakers increase their voice level under DAF (e.g., Fairbanks, 1955). The louder a speaker speaks, the louder (and more intrusive) the delayed signal would be. Such effects would be strategic, rather than dependent upon feedback. Howell et al. found that subjects who raised their voices a lot under DAF were more disrupted

than speakers who did not raise their voices much. Thus, susceptibility can be explained on the basis of this hypothesis. Similar factors may account for differences across the sexes and developmentally in susceptibility to the effects of DAF.

Finally, the hypothesis predicts that the proven usefulness of DAF in the treatment of stuttering is unlikely to work, because the technique corrects something that is deviant in the feedback of such speakers (see, e.g., Leith & Chmiel, 1980, and Webster & Lubker, 1968). If the effects of DAF in normals are not to do with feedback regulation, then they should not be with stutterers either. Two proposals that have been made about abnormal feedback in the stutterer are that they have unusual middle-ear muscle activity which affects their feedback (Webster & Lubker, 1968) and that their bone-conducted feedback is deviant (Cherry & Sayers, 1956). It appears that certain of the effects of speaking under altered feedback can be explained by psychoacoustic effects associated with the altered feedback (Howell & Powell, in press).

Table 4
Estimates of Delay (in Milliseconds) at Which Peak Occurred for the Various F1 and F2/3 Levels in the Part Delay Conditions of Experiment 2

		F1 Level		
		0	+10	+20
	Group 2a			
	0	n.s.	250	250
F2/3 Level	+10	n.s.	n.s.	215
	+20	n.s.	250	215
	Group 2b			
	0	n.s.	140	n.s.
F2/3 Level	+10	n.s.	180	n.s.
	+20	280	260	280

Note—The delay at which a peak occurred was only estimated in conditions where regression analysis indicated that a quadratic component was needed to provide a good fit.

REFERENCES

BLACK, J. W. The effect of delayed side tone upon vocal rate and intensity. Journal of Speech and Hearing Disorders, 1951, 16, 56-60.

BORDEN, G. J. An interpretation of research on feedback interruption in speech. Brain and Language, 1979, 7, 307-319.

BUTLER, R. A., & GALLOWAY, F. T. Factoral analysis of the delayed speech feedback phenomenon. Journal of the Acoustical Society of America, 1957, 29, 632-635.

CHERRY, C., & SAYERS, B. Experiments upon the total inhibition of stammering by external control, and some clinical results. Journal of Psychosomatic Research, 1956, 1, 233-246.

ELMAN, J. L. Effects of frequency-shifted feedback on the pitch of vocal productions. Journal of the Acoustical Society of America, 1981, 70, 45-50.

FAIRBANKS, G. Selected vocal effects of delayed auditory feedback. Journal of Speech and Hearing Disorders, 1955, 20, 333-345.

FANT, G. M. C. Acoustic analysis and synthesis of speech with applications to Swedish. Ericsson Technics, 1959, 1, 1-105.

HOWELL, P., & POWELL, D. J. Hearing your voice through bone and air; Implications for explanations of stuttering behavior

from studies of normal speakers. *Journal of Fluency Disorders*, in press.

HOWELL, P., POWELL, D. J., & KHAN, I. Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 1983, 9, 772-784.

LANE, H. L., & TRANEL, B. The Lombard Sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, 1971, 14, 677-709.

LEE, B. S. Some effects of side tone delay. *Journal of the Acoustical Society of America*, 1950, 22, 639-640.

LEITH, W. R., & CHMIEL, C. C. Delayed auditory feedback and stuttering: Theoretical and clinical implications. In N. J. Lass (Ed.), *Speech and Language: Advances in basic research and practice* (Vol. 3). London: Academic Press, 1980.

SIEGEL, G. M., & PICK, H. L. Auditory feedback in the regulation of the voice. *Journal of the Acoustical Society of America*, 1974, 56, 1618-1624.

WEBSTER, R. L., & LUBKER, B. Interrelationships among fluency-producing variables in stuttered speech. *Journal of Speech and Hearing Research*, 1968, 11, 754-766.

YATES, A. J. Delayed auditory feedback. *Psychological Bulletin*, 1963, 60, 213-232.

NOTES

1. As a check on whether the filter conditions made the vowels more difficult to identify, the following experiment was run. Six subjects were instructed on what the seven vowels /i/, /I/, /ɛ/, /a/, /u/, /ʌ/, and /ɔ/ sounded like and given orthographic symbols to represent each of the vowels. These were to cover a range of vowel qualities and to serve as responses. The subjects were permitted to use their own additional symbols in responding and to explain them to the experimenter after completion of the experiment. A set of six lists, each comprising 30 vowels (10 /a/s, 10 /u/s, and 10 /i/s) in random order was recorded by a male speaker (1 sec between vowels). The subjects listened to two lists each under the conditions of normal speech, speech low-pass fil-

Table A1
Error Rate Across Filter Conditions for the Vowels
/a/, /I/, and /u/ for the Experiment
Described in Footnote 1

	/i/	/a/	/u/
Normal	11.7%	5%	11.7%
Low-pass Filter	51.7%	64.2%	79.2%
Bandpass Filter	40.8%	35.8%	38.3%

tered at 550 Hz, and speech bandpass filtered at between 550 Hz and 3.5 kHz (all filter roll-offs 48 dB/octave). Two lists were given in each condition to counterbalance for order effects. The listeners had to write the symbol corresponding to the vowel they thought they had heard. The number of errors was calculated for each vowel under each filter condition. These are presented for each vowel and each filter condition in Table A1.

An analysis of variance was performed to ascertain whether the filtering conditions increased error rate and whether this differed across vowels. The only effect was a difference across filter conditions [$F(2,10) = 32.1$, $p < .001$]. A Newman-Keuls test showed that the differences between the normal speech condition and the two filter conditions and the difference between the bandpass and low-pass filter conditions were significant (all $p < .01$). Inspection of the table shows highest error rate in the low-pass filter condition, followed by the bandpass filter condition, and least errors in the normal speech condition. This indicates that filtering affects subjects' ability to identify vowels.

2. Although the filter cutoff of the low-pass filter (550 Hz) is lower than measurements of F1 for /a/ (roughly 650 Hz for these speakers), the peak value of F1 would be attenuated by only about 12 dB. Some attenuation of F1 was tolerated in this condition to get a good separation of F1 from the higher formants.