

Differential effects of two nonspeech sounds on phonemic restoration

BARRY LAYTON

Departments of Psychology and Neurology, St. Louis University, St. Louis, Missouri 63103

Sixty-nine introductory psychology students participated in a phonemic restoration (PHR) experiment. PHR is the illusory perception of a phoneme deleted from a spoken sentence when that phoneme is replaced by an extraneous sound. Contrary to previous results, it was found that PHRs were significantly more prevalent when a cough was substituted, than when a tone replaced the deleted speech sound ($p < .05$). The study also revealed a decreasing tendency across trials to mislocalize the extraneous cough or tone prior to its veridical location when expectancy of stimulus position was minimized. The theoretical import of this latter finding is discussed, as is a methodological problem common to all PHR research and its possible effect on the percentage of PHRs reported.

Warren (1970) reported that the deletion of a phoneme from a tape-recorded sentence could not be detected if it was replaced by a cough or a tone. The term "phonemic restoration" (PHR) was coined to describe illusory perception of the missing speech sound. PHRs did not occur when a silent gap was inserted in place of the missing phoneme.

In a follow-up study, Warren and Obusek (1971) replicated and extended the earlier findings. The second study confirmed the existence of PHRs when sounds other than a cough or tone replaced the missing phoneme. The use of a repeated measurement design in the second study also permitted the discovery that the perceived position of the intrusive sound occurred prior to its veridical position on the first hearing of the tape, and that on subsequent hearings a shift in direction towards its veridical position occurred. In addition, Warren and Obusek found that PHR and mislocalization occurred even in the case when subjects could (by monitoring a typewritten statement of the stimulus sentence while listening to it) make the position judgment simultaneously with stimulus presentation.

The present study tested the effects of altering the repeated measurement design used by Warren and Obusek. In the Warren and Obusek design, the stimuli were apparently presented in such a way that subjects were aware that they were listening to the same recording four successive times. In the present experiment, subjects were unable to determine whether the four successive presentations were identical with respect to the location of the intrusive stimulus and with respect to whether the intrusion completely replaced the phoneme. The procedural

variation was introduced because it was felt that the temporal shift in localization found by Warren and Obusek might not occur under conditions in which each judgment was more or less independent of the preceding one.

METHOD

Subjects

Thirty-eight male and 31 female students from the introductory psychology subject pool at St. Louis University volunteered to participate in the study. Although experimental participation was not a course requirement, subjects were compensated by means of course credit.

Stimulus Materials

A speaker, selected both for clarity of voice and for minimal regionally specific dialect, recorded the sentence, "The state governors met with their respective legislatures convening in the capital city." The duration of the recorded sentence was 4.95 sec, well within the limits of average duration of a single spoken sentence of 25 syllables (Lenneberg, 1967, Chap. 3). The recording was made on Soundcraft S-6 acetate recording tape at a speed of $7\frac{1}{2}$ in./sec on a Sony TC-630 stereophonic tape recorder. A copy of the recording was made onto an identical recorder using Scotch Mylar recording tape.

The portion of the original acetate recording containing the first "s" in the word "legislatures," along with transitional portions of the preceding "i" and following "l," was excised in the manner described by Warren and Obusek (1971). The excised section measured 3.8 cm (200 msec) in length. A section of used blank tape, also 3.8 cm in length, was spliced into the master in place of the deleted tape. Subsequent comparison of copies of the original and edited tapes on a Brüel and Kjaer Model 2112 spectrometer and Model 2305 level recorder confirmed that the first "s" in "legislatures" was entirely absent from the edited version.

The three stimulus conditions, in which a cough, tone, or silent gap replaced the deleted phoneme, were produced in the manner suggested by Warren and Obusek. The silent-gap condition was produced by copying the edited master four successive times, using the two Sony recorders, onto the Scotch Mylar tape at $7\frac{1}{2}$ in./sec. A 30-sec interval was maintained between each copy. A recording of a 1,000-Hz tone was then made. Oscilloscopic comparisons of the edited master with the tone showed that the tone produced was 6 dB above the peak intensity of the speech. A recording of the cough, of peak intensity equal to that of the tone, was also made. Sections of the cough and tone recordings 200 msec in length were

The author gratefully acknowledges the assistance at various stages of the project of James Hauschultz, Daniel O'Connell, and James Mulligan. This paper is sponsored by D. Gene Davenport, who takes full editorial responsibility for its contents.

excised. The splice in the edited master was opened; first the tone and then the cough sections were inserted in place of the blank tape. Copies of these newly synthesized recordings were produced in the same manner as described for the silent-gap version.

Procedure

Subjects were randomly assigned to the cough, tone, or silent-gap condition. Subjects were tested individually, seated at a table facing a blank wall. After being seated, subjects read the following instructions (modified from Warren and Obusek): "This is an experiment in auditory perception. You will hear a sentence spoken clearly. A cough (silent gap, tone) will occur at some time during the sentence. You are to indicate where you think the cough (silent gap, tone) occurred by placing a circle around the EXACT place in a typewritten statement of the sentence. You are to indicate also whether the cough (silent gap, tone) completely replaced the sound(s) which you circle. This procedure will be repeated several times, with a short delay between each presentation."

The experimenter offered to answer any possible questions. When it was clear that subjects understood the task, the testing procedure began.

Stimuli were presented through a pair of Telex Combo headphones rewired for monophonic delivery. The peak sound level of the stimulus sentence was 80.5 dB/.0002 μ bar in the left earphone and 79.5 dB/.0002 μ bar in the right earphone. These parameters, which closely matched those used by Warren, were confirmed with a B & K Type 2203 sound-level meter, Type 4152 artificial ear, Type 4132 microphone, and Type 1613 octave filter set positioned at 1,000 Hz. The tape recorder used for stimulus delivery was a Wollensak 3M monophonic Model 6020 AV.

During the testing procedure, the experimenter stood behind the subject, monitoring the stimulus presentation by observing color-coded leads and trailers previously attached before and after each copy of the recording. For the first three trials, immediately following stimulus presentation, the experimenter handed subjects an answer sheet consisting of a typewritten statement of the stimulus sentence and a two-choice question in which the subject was to indicate whether or not the extraneous sound or silence had completely replaced the circled speech sound(s). The answer sheet was withdrawn prior to each subsequent trial. For the fourth trial, the answer sheet was handed to subjects prior to the stimulus presentation. The method of stimulus presentation was to keep the recorder turning until all four copies of the sentence were heard. Using this technique, subjects could not assume that the sentences were identical with respect to the position of the extraneous stimulus and whether or not the phoneme absent on one trial would be absent on succeeding trials.

RESULTS

Nine of the 69 subjects were not included in the data analysis. It was decided as part of the procedure that the first subject in each condition would serve only to familiarize the experimenter with the running

Table 1
Absolute Temporal Error Magnitudes in Phonemic Units by Groups of 20 Subjects, Each Making Four Successive Judgments

Task	Medians, First and Third Quartiles			
	Trial 1	Trial 2	Trial 3	Trial 4
Silent	1.91	1.55	1.08	1.55
Gap	(.75, 3.25)	(.75, 2.13)	(.375, 1.75)	(.63, 1.95)
Cough	4.75	1.25	1.65	1.25
	(1.25, 23.75)	(.89, 3.25)	(1.00, 7.75)	(.75, 1.91)
Tone	3.58	1.65	1.50	1.18
	(.75, 19.50)	(1.08, 5.75)	(1.05, 2.25)	(.82, 2.91)

Table 2
Directional Temporal Error Magnitudes in Phonemic Units by Groups of 20 Subjects, Each Making Four Successive Judgments

Task	Medians, First and Third Quartiles			
	Trial 1	Trial 2	Trial 3	Trial 4
Silent	-1.50	-1.35	-.75	-1.45
Gap	(-2.75, +.25)	(-1.92, -.06)	(-1.75, .0)	(-1.95, -.42)
	-3.25	+.25	-.25	+.25
Cough	(-19.75, -.75)	(-1.75, +1.08)	(-2.25, +1.42)	(-1.54, +1.15)
	-2.75	-1.50	-1.13	-.75
Tone	(-11.75, +.25)	(-3.75, -.75)	(-2.00, +.75)	(-1.42, +.91)

of the experiment. Six other subjects were eliminated: two had a previous history of hearing disorders, three failed to complete the answer sheets as per instructions, and one was inadvertently handed the improper instructions. The data were analyzed on the basis of 20 subjects in each condition.

There were two dependent variables of interest. One was the mislocalization score, or error of localization in phonemic units from the origin of the actual location of the extraneous stimulus. The second was the qualitative judgment of whether or not the extraneous cough, tone, or silent gap completely replaced the speech sound(s) circled on the answer sheet.

In the analysis of mislocalization, the occurrence of several extreme scores dictated the median as the measure of choice. For the same reason, both Warren (1970) and Warren and Obusek (1971) also used this measure. For the determination of localization errors, each phoneme and each space between words was counted as one unit. When more than one phoneme was circled, the mean error was recorded.

Table 1 presents the absolute error scores for all four trials of each condition. These scores do not reflect the direction of the error, only its magnitude.

For the first trial, which represented a replication of the Warren (1970) experiment, the difference in absolute error scores was not significant between the silence and the tone conditions or between the cough and tone conditions (Mann-Whitney U test). However, the magnitude of error in localizing the cough was significantly greater than in localizing the silence ($p < .05$). None of these three comparisons were significant on the second trial, although, on the third trial, the silent gap was localized significantly more accurately than both the cough ($p < .01$) and the tone ($p < .05$). On the fourth trial, when subjects had the answer sheet during stimulus presentation, there were no significant differences.

In Table 2, the direction of the mislocalization error was taken into consideration.

On the first trial, mislocalizations for all three conditions were earlier than the veridical position of

the extraneous stimulus. For each condition, a shift to a relatively later position can be observed. This shift was significant (Wilcoxon signed rank test) only in the tone ($p < .01$) and in the cough ($p < .001$) conditions. This analysis included shifts from Trials 1 through 3 only, Trial 4 consisting of a different method of presentation.

In terms of the occurrence of the PHR illusion, the indicator used by Warren and by Warren and Obusek was the incidence of subjects reporting all phonemes present. This indicator was also used in the present study. Obviously, a subject reporting all phonemes present cannot have perceived the absence of a phoneme. These data are presented by groups in Table 3.

The Fischer exact probability test was used to determine if there was a difference between the groups in terms of PHR. No comparison was significant for the first trial. For the second trial, the cough condition showed significantly more PHRs than both the silent gap condition ($p < .01$) and the tone condition ($p < .05$). On the third trial, the cough group again reported significantly more PHRs than both the tone and the silence groups ($p < .05$). There was no significant difference on the fourth trial.

Other measures inferring the occurrence or nonoccurrence of PHRs are possible. One such measure is the percentage of subjects judging that the extraneous cough, tone, or silence replaced a speech sound *and* included the target phoneme in their judgment. It was assumed that subjects meeting this criterion did not produce PHRs. The lower the percentage of subjects meeting this criterion, the greater the occurrence of PHRs. Table 4 presents this data.

For the first trial, the Fischer exact test showed that only the difference between the cough and tone groups was significant ($p < .05$) using the second criterion. No comparisons were significant for the second trial. On Trial 3, the group hearing the cough had significantly more PHRs than both the silence and tone groups ($p < .05$). Again, on Trial 4, there were no significant differences.

DISCUSSION

Stimulus presentation on Trial 1 of the present experiment constituted a replication of Warren's (1970) experiment. The results of this replication were equivocal with respect to confirmation of the earlier findings. Mislocalization of the tone and cough was approximately of the same magnitude as that reported

Table 4
Percentage of Subjects Judging Both that the Extraneous Event Replaced a Speech Sound *and* Who Included the Target Phoneme in Their Judgment

Task	Trial 1	Trial 2	Trial 3	Trial 4
Silent Gap	25	30	45	30
Cough	5	20	10	40
Tone	35	35	40	25

by Warren. (Warren did not report the magnitude of mislocalization in the group receiving the silent gap.) In terms of PHRs, for subjects hearing the cough, the present study found support for the Warren findings. Sixty-five percent of the subjects in the cough group reported all speech sounds present, and 95% were unable to identify the target phoneme as being absent. On the other hand, the tone apparently was not effective in the production of PHRs. Only 40% of the subjects in that group reported that all speech sounds were present, and 35% could include the target phoneme as one of those believed to have been completely replaced by the tone. Since only 25% of the subjects in the silence condition could make the same determination, the tone seems to have been less effective than the silence (which Warren found to be ineffective) in the production of PHRs. These results are inconsistent with Warren's finding that 100% of the subjects in his tone group reported all speech sounds present. Warren and Obusek also found almost universal susceptibility to the illusion across four presentation trials. The inconsistency of the tone effect is paradoxical because subjects in the present study had *greater* difficulty in locating the tone's position than did Warren's subjects.

The paradox is more apparent than real, however. The mislocalization phenomenon is independent of the occurrence of the illusion. Ladefoged and Broadbent (1960) reported that a brief nonspeech sound could not be located within an intact sentence (i.e., one without deletion). This independence logically explains why the mislocalization effect may occur in the absence of the illusion.

The inability to replicate the PHR illusion in the tone condition might be explained by the fact that the deletion in the tape Warren (and Warren and Obusek) used was 120 msec in duration, while the deletion in the present study was 200 msec. Although both durations represent average values in terms of normative language data for spoken phonemes, it is reasonable to suppose that the longer deletion might be more easily perceived than the shorter one. The fact that the missing phoneme was "heard" in the cough condition (with the exception of one subject) might be due to the fact that the cough, being a familiar source of speech interference, is a more effective aid in the production of PHRs than is the tone. The latter is an artificial laboratory speech masker not occurring in the daily environment. The relative strength of the cough effect may be so compelling that a difference of 80 msec would not suffice to weaken its effect. The masking effect of the tone, not experienced prior to the experimental situation, may be more dependent on its duration, to the extent that its effect would be functionally equivalent to that of the cough at short durations and functionally equivalent to the silent gap at longer durations.

The present study involved a variation of the procedure employed by Warren and Obusek. These investigators found a decreasing tendency across trials to perceive the extraneous stimulus prior to its veridical position. Because they did not specify the method of successive presentation used, it was assumed that the same stimulus tape was replayed so that subjects were aware that they were hearing the identical recording on successive trials. In such circumstances, the subject is permitted the opportunity of anticipating the general area of localization, possibly enabling him to make finer judgments from trial to trial. Change in the direction of the error might therefore be attributable to increasing familiarity with the position of the extraneous cough, tone, or silent gap, rather than to familiarity with the meaningfulness of the sentence as Warren and Obusek have argued. In other words, a timing skill

Table 3
Percentage by Groups of Subjects Judging that the Extraneous Event Did Not Replace Any Speech Sound

Task	Trial 1	Trial 2	Trial 3	Trial 4
Silent Gap	40	40	40	40
Cough	65	80	70	55
Tone	40	50	40	40

(involving learning) rather than a language skill could account for the shift. In the present study, four copies of the sentence were made in such a way that the tape kept on rolling throughout the experiment. Playback of the tape was unnecessary and subjects could not possibly know whether each presentation was identical to the preceding ones with respect to position of the extraneous stimulus and with respect to the deletion of a phoneme. Judgments of each type were thereby rendered more or less independent of the preceding presentations. The present experiment produced the "left-to-center" shift in localization found by Warren and Obusek. Apparently, then, the shift is not dependent on the anticipation of identical stimulus presentations. However, the fact that the absolute difference in mislocalization scores for the three conditions tended to equal out across trials and that the difference in error between groups was not significant on the fourth trial (when subjects could follow the recorded sentence with the typewritten statement) indicates that the magnitude of the mislocalization phenomenon in all conditions was reduced across trials in the direction of the magnitude originally observed for the silent gap. This much of an equalization was not observed in Warren and Obusek's study, and might, therefore, be attributable to the difference in procedure.

A final note concerning the instructions is in order. For reasons of continuity, the instructions used in the present study were identical to those used by Warren and Obusek (except for the addition by the present author of the first and last sentences). An examination of these instructions reveals an unfortunate ambiguity. The instructions ask for a judgment as to whether the circled sounds (indicating the perceived position of the nonspeech sounds) were completely replaced by the extraneous cough, tone, or gap. Warren and Obusek assumed (as did the present author) that, if the subjects circled more than one phoneme and checked the box indicating that they were not completely replaced by the extraneous sounds, all of the circled phonemes were perceived. This assumption was, in fact, the basis for determining the existence or

nonexistence of PHRs, since virtually every subject in the present experiment circled three or more phonemes. Unfortunately, this assumption is highly questionable. Warren and Obusek reported that subjects are confused about the position of the extraneous stimulus: their inability to locate the cough or tone is not the result of an illusory perception of its location, but is due to a conscious confusion about its location. Subjects were generally aware that their positional judgment was a guess rather than a positive determination. A conservative guesser might circle three or four phonemes, feeling only that the nonspeech stimulus occurred somewhere in the area, and knowing that the stimulus did not extend the whole length of the circled area. In this case, he would reply that the cough, tone, or gap did not replace the circled speech sounds if he heard any *one* of them. Obviously, then, this response does not necessarily imply a PHR. The possibility that the instructions may be interpreted in the manner outlined above would result in a report of PHRs higher than their true incidence. Future experiments in the area should modify the instructions so as to eliminate the confounding ambiguity.

REFERENCES

- LADEFOGED, F., & BROADBENT, D. E. Perception of sequence in auditory events. *Quarterly Journal of Experimental Psychology*, 1960, **12**, 162-170.
- LENNEBERG, E. H. *Biological foundations of language*. New York: Wiley, 1967.
- WARREN, R. M. Perceptual restoration of missing speech sounds. *Science*, 1970, **167**, 392-393.
- WARREN, R. M., & OBUSEK, C. J. Speech perception and phonemic restorations. *Perception & Psychophysics*, 1971, **9**, 358-362.

(Received for publication July 23, 1975.)