Journal of Experimental Psychology
1970, Vol. 85, No. 3, 411–417

PREPERCEPTUAL AUDITORY IMAGES [1]

DOMINIC W. MASSARO [2]

University of California, San Diego

The *S*s were required to identify the pitch of a 20-msec. test tone. In four experiments, the test tone was followed by a masking tone after a variable silent intertone interval. The masking tone interfered with the perceptual processing of the test tone such that pitch identification performance improved with increases in the silent intertone interval. The results also indicated that (*a*) the amount of interference produced by the masking tone was relatively independent of the similarity of the test and masking tones, (*b*) dichotic (contralateral) masking was as effective as the binaural case, and (*c*) presenting the masking tone before the test tone did not disrupt pitch-identification performance. The experiments are interpreted as a demonstration of the existence of a central auditory image that remains after a short tone burst is terminated.

Extant theories of perception and memory usually assume the existence of a sensory information-storage system that retains primitive physical features of a stimulus presentation even after the stimulus is terminated (Neisser, 1967; Norman, 1968). Since the processing of information usually requires the analysis of a temporal continuum of sensory features, some sensory storage of the early impressions during the later inputs is necessary for perception of the meaning of the message. A number of studies have provided evidence for a visual image that persists after the presentation of a visual stimulus (Averbach & Coriell, 1961; Posner & Keele, 1967; Sperling, 1960; Weisstein & Haber, 1965). However, very little direct evidence exists for an auditory information store analogous to the visual store. The present experiments provide evidence for an acoustic storage system that retains an auditory image after a short tone is terminated.

Perception or perceptual processing can be thought of as an analysis of the sensory input in which features are detected that correspond to encoded features in memory (Massaro, in press). When a sufficient number of features are found in the sensory input that correspond to the features of the item in memory, the stimulus is recognized as that item. Since the feature detection takes time, a sensory image could hold the sensory experience after the stimulus is terminated. Accordingly, the time course of the image can be determined by terminating the feature-detection process at different times after a short stimulus is presented. In the present study, a short high or low frequency tone was presented and *S*s identified the pitch of the test tone as high or low. Given a sufficient frequency difference between the high and low tones, *S*s have no difficulty performing this task. However, if the test tone is followed immediately by a retroactive (masking) tone, identification performance decreases markedly. The purpose of the present study is to demonstrate that this decrease in performance is due to the masking tone's interference with the perceptual processing of the auditory image of the test tone. Furthermore, varying the silent interval between the test and masking tones should indicate the temporal course of perceptual processing of the sensory image of the test tone.

EXPERIMENT I

Method

Subjects.—Two females and one male attending the University of California were employed in the present study. They were paid $1.88 an hr. for their services.

[1] This investigation was supported in part by a National Institutes of Health Postdoctoral Fellowship (MH 39369-02) from the United States Public Health Service.

[2] Requests for reprints should be sent to Dominic W. Massaro, who is now at the Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

Procedure.—In the present experiment, *S*s were required to identify the pitch of a test tone. The duration of the test tone was 20 msec. On any trial, one of two tones could be presented. The two test tones were programmed to occur equally often. The *S*'s task was to identify the higher tone (870 Hz.) as "high" and the lower tone (770 Hz.) as "low." A second tone (820 Hz.), referred to as the masking tone, followed the test tone after a variable silent intertone interval. The masking tone lasted 500 msec. The silent intertone interval lasted 0, 20, 40, 80, 160, 250, 350, or 500 msec. The loudness of the test and masking tones was 81 db.

The *S*s were tested simultaneously in sound-insulated chambers (Industrial Acoustics). All experimental events were controlled by a PDP-8 computer. A digitally controlled oscillator (Wavetek Model 155) was used to produce the pure tones. The stimuli were presented binaurally over matched headphones (Grason-Stadler Model TDH-39). Each *S* recorded his "high" or "low" decision by pressing one of two pushbuttons labeled "high" and "low," respectively. Following the 2-sec. response period, feedback was given by illuminating a small light for 500 msec. above the correct response button. The intertrial interval was 2½ sec.

On every trial, *S*s heard a test tone followed by a variable silent interval followed by the masking tone. They identified the test tone as "high" or "low" and were then informed of the correct answer for that trial. The *S*s were practiced in this task for 2 wk. (about 15 hr.) before the present study. The first few practice sessions were given without the masking tone present so that *S*s could learn the high and low tones. In the present experiment, all 16 experimental conditions were completely random within a given session. A session consisted of 400 trials. The results consist of the data of two sessions per day for 4 days giving about 200 observations at each experimental condition for each *S*.

Results

Figure 1 presents the percentage of correct identifications of the test tone for each *S* as a function of the duration of the silent interval before the onset of the masking tone. Since the percentages are determined from both the high and low tone trials, about 400 observations contribute to each data point. The results indicate that identification performance increased as the intertone interval increased up to about 250 msec. Therefore, these *S*s were able to continue processing the auditory image of the test tone during the silent retroactive interval and this processing improved identification performance. That is, the feature detection necessary for identification of the pitch of the test tone took place during the test-tone presentation and the silent interval afterwards, but was interfered with during the masking tone presentation. The results might be interpreted as indicating that the auditory image of the test tone lasted about 250 msec. However, as noted in the Discussion section, the apparent duration of the image will probably be dependent upon the nature of the perceptual task.

Another explanation of the results of the present study might be that the masking tone prevents the sensation of the test tone from taking place. Given that the masking tone was of the same loudness as the test tone, it is unlikely that the masking tone overtook the test tone in the auditory pathways and prevented the sensation. The *S*s reported that they heard the test tone on every trial even though they were not always able to identify its pitch. A second experiment was carried out to provide additional evidence for the fact that the results of the first study do not represent masking of the sensation of the test tone.

Experiment II

Method

Subjects.—The *S*s employed in Exp. I were used in this study. Experiment II started the day after Exp. I ended.

Procedure.—The procedure was identical to that of Exp. I with the exception that the masking tone was presented before the test tone. Therefore, the presentation of the masking tone lasting 500 msec. started a trial. After a variable silent intertone interval of 0, 20, 40, 80, 160, 250, 350, and 500 msec., the test tone was presented for 20 msec. The frequencies and amplitude of the test and masking tones were the same as in Exp. I. On every trial, *S*s heard the masking tone followed by the test tone. They identified the test tone as "high" or "low" and were then informed of the correct answer for that trial. All experimental conditions were completely random within a given session. Due to the results of the first day (two sessions of 400 trials each), the experiment was terminated at this point.

Results

Table 1 presents the proportions of correct identifications as a function of the intertone interval for each of three *S*s. The results indicate that *S*s correctly identified the test tone at least 94% of the time at each intertone interval. If the masking tone was interfering with the sensation of the test tone, it should do so equally whether presented before or after the test tone (Homick, Elfner, & Bothe, 1969; Kahneman, 1968). Thus, the results support the hypothesis that a retroactive masking tone can interfere with the perceptual processing of the auditory image without necessarily disrupting the initial sensory reception of the tone.

Another study was carried out to determine if processing of the auditory image can be disrupted by a masking tone that is presented to the ear contralateral to the ear of the test-tone presentation. This result might indicate that the auditory image is not represented in the cochlea, but resides centrally. Given a central auditory information store, contralateral interference of the image should not be surprising since the projections of the two ears are completely intermingled in the auditory cortex (Geldard, 1953; Netter, 1967).

Experiment III

Method

Subjects.—Two of the *S*s in Exp. I and II were employed in the present study.

Procedure.—Experiment III replicated Exp. I except that a dichotic masker was employed. Therefore, the exact procedure of Exp. I was used in the present study with the exception that the test tone was presented to one ear and the masking tone was presented to the opposite ear. Within a given session, the test tone was always presented to the same ear followed by the masking tone presented to the opposite ear. The experiment lasted 4 days with two sessions of 400 trials each per day.

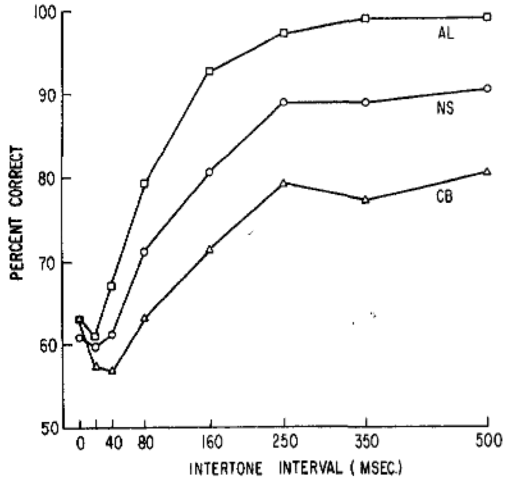

Fig. 1. Percentage of correct identifications of the test tone for *S*s AL, NS, and CB as a function of the duration of the silent intertone interval in Exp. I.

Results

Figure 2 presents the percentage of correct identifications for both *S*s as a function of the intertone interval between the test and masking tones. The results show an overall performance decrement with respect to Exp. I (cf. Fig. 1). There was no improvement in identification in the first 80 and 160 msec. of the silent interval for *S*s NS and CB, respectively. However, performance did improve with further increases in the silent interval for both *S*s. Thus, Exp. III provides some evidence that a retroactive masking tone presented to the contralateral ear of the test tone interferes with the perceptual processing of the test tone. Since only two *S*s were employed in the present experiment, another study was carried out employing a contralateral mask-

TABLE 1

Proportions of Correct Identifications of the Test Tone for Each *S* as a Function of the Silent Intertone Interval before the Test Tone Presentation in Experiment II

S	Intertone interval (in msec.)							
	0	20	40	80	160	250	350	500
AL	1.0	1.0	1.0	1.0	1.0	.983	1.0	1.0
NS	.943	.964	.987	.982	.985	.969	.975	1.0
CB	.985	.986	.985	.981	1.0	1.0	1.0	.988

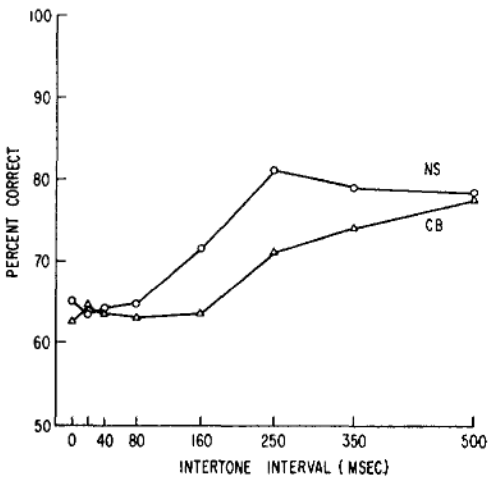

FIG. 2. Percentage of correct identifications of the test tone for *S*s NS and CB as a function of the duration of the silent intertone interval in Exp. III.

ing tone. In addition, the frequency of the masking tone was varied to provide evidence concerning the masking effects of a tone that is somewhat dissimilar from the test tone.

EXPERIMENT IV

Method

Subjects.—Three new *S*s were employed in the present experiment. The two females and one male were students at the University of California and were paid $1.88 an hr.

Procedure.—Experiment IV replicated Exp. III in all but two respects. First, the test tone was presented to either ear in a random fashion so that the test tone was equally likely to occur in the left or right ear on any trial. The masking tone was presented to the ear contralateral to the test-tone presentation. Second, the frequency of the masking tone was either 820 or 999 Hz. One session was presented per day at each of the two masking tone frequencies. All other experimental conditions were completely random in a given session. The experiment lasted 5 days with two sessions of 400 trials each per day. The *S*s had 3 days of practice (six sessions) before the present experiment.

Results

Figure 3 presents the percentage of correct identifications for each *S* as a function of the frequency of the masking tone and the silent intertone interval. The figure shows that, overall, identification performance improved as the duration of the silent interval increased. An analysis of variance of *S*s, masking tone frequencies, and intertone intervals performed on the percentage of correct identifications revealed a significant effect of the intertone interval, $F\ (7,\ 14) = 11.81$, $p < .001$. Contrary to expectations, the dissimilar (999 Hz.) masker produced more overall interference with pitch identification than did the similar (820 Hz.) masker, $F\ (1,\ 2) = 65.66$, $p < .001$. However, Fig. 3 shows that this relationship only existed for two *S*s at intertone intervals of 80 msec. or larger.

The present results replicated those of Exp. III by showing that a retroactive masking tone presented contralateral to the test tone can disrupt identification of the test

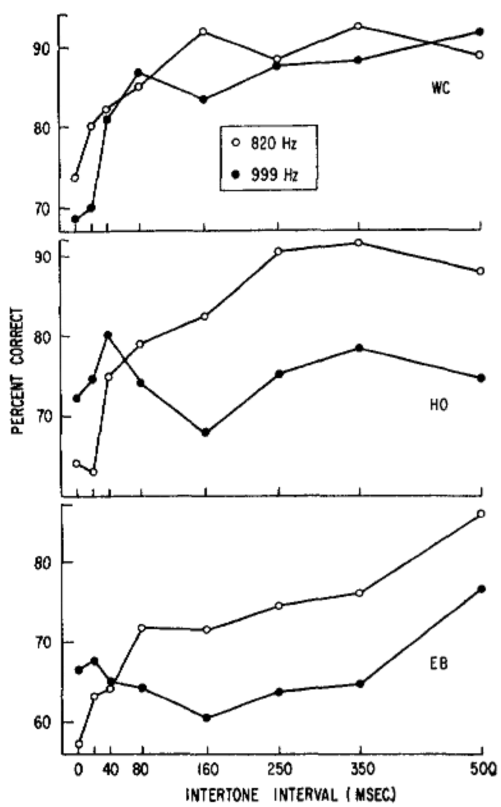

FIG. 3. Percentage of correct identifications of the test tone for *S*s WC, HO, and EB as a function of the frequency of the contralateral masking tone and the silent intertone interval in Exp. IV.

tone. The results support the idea of a transient auditory image that remains after a short tone is presented. Since performance improves with increases in the silent intertone interval, some sensory experience must remain so that it can be processed in order to identify the stimulus that produced the experience. Also, given that the sensory image is located centrally, a dichotic masker would be as effective as the binaural case. The results of Exp. III and IV, therefore, support the hypothesis of a central auditory information store.

Figure 3 indicates that the similarity of the masking tone to the test tone may be an important factor in the nature of the interference effects of the masking tone. Accordingly, an experiment was carried out specifically to analyze the effects of similarity on interference of the processing of the image of the test tone. In Exp. IV, the frequency of the masking tone was held constant within a given session. The *S*s could, therefore, employ different strategies for the two masking tone conditions. This could account for the discrepancy found between the two masking conditions for *S*s HO and EB (cf. Fig. 3). To avoid this possibility, the frequency of the masking tone was varied randomly from trial to trial in Exp. V.

Experiment V

Method

Subjects.—Two females and one male were employed in the present study. The *S*s WC and EB also participated in Exp. IV.

Procedure.—The procedure was essentially the same as Exp. I except that different frequencies of the test and masking tones were employed. In the present task, the low and high tones were 480 and 600 Hz., respectively. The frequency of the masking tone was varied randomly from trial to trial. The four frequencies were 540, 660, 720, and 999 Hz. These test and masking tone frequencies have been shown to be perceived as equally loud at the 81-db. signal level used in the present experiment (Fletcher & Munson, 1933). The durations of the test tone, masking tone, and silent intertone intervals were identical to those in Exp. I. All experimental conditions were completely random within a given session. The results consist of the data of two sessions per day for 10 days. The *S*s had 3 days of practice in this task before the present experiment.

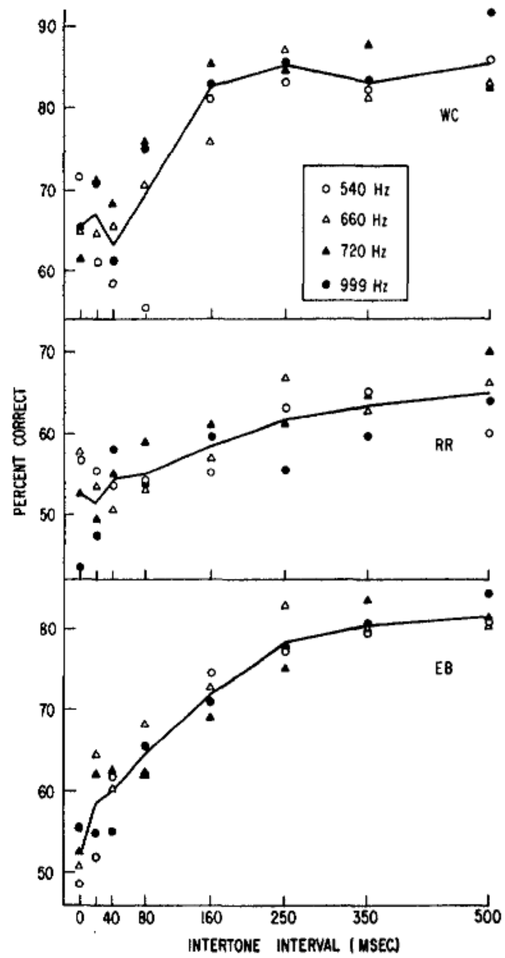

FIG. 4. Percentage of correct identifications of the test tone for *S*s WC, RR, and EB as a function of the frequency of the masking tone and the silent intertone interval in Exp. V. The continuous lines are means averaged over masking tone frequency.

Results

Figure 4 presents the percentage of correct identifications for each *S* as a function of the frequency of the masking tone and the silent intertone interval. The figure shows no large differences in performance as a function of the frequency of the masking tone. Furthermore, the shape of the masking function appears very similar under the different masking tone frequencies. To test these hypotheses, an analysis of variance of *S*s, masking tone frequencies, and intertone intervals was performed on the percentage of correct identifications. The analysis indicated no differences in overall performance as a function of the masking tone frequencies, F (3, 6) = .97, and a nonsignificant effect of the interaction of the masking tone frequency and the intertone interval, F (21, 42) = .94. As expected, the improvement in performance as a function of the silent intertone interval was significant, F (7, 14) = 15.61, $p < .001$. These results indicate that the masking tone can be relatively dissimilar from the test tone and still interfere with its pitch identification.

Assuming that both perceptual and decision processes operate in an identification task, the present experiments were concerned with the perceptual system. However, to test whether the nature of the results were influenced by decision factors (e.g., response biases), each experiment was also analyzed in terms of strength measures analogous to the value d' of the theory of signal detectability (Green & Swets, 1966). This measure is assumed to be independent of response biases and has been shown to provide an orderly measure of the temporal course of perception and memory (Massaro, in press; Wickelgren & Norman, 1966). The analyses of the percentage of correct identifications and the strength values revealed no differences in results and the percentages were employed simply for ease of exposition.

Discussion

The present studies have shown that the pitch identification of a short test tone can be interfered with by a retroactive masking tone. Therefore, pitch identification improved as the silent interval between the test and masking tones increased. These results are interpreted as demonstrating the existence and utility of an auditory information storage system that retains a primitive physical representation of auditory stimulus after it is terminated. Accordingly, in the present study, the image of the test tone was processed during the silent interval in order to identify the tone's pitch.

Three results also provide information concerning the nature of the interference effects of the masking tone. First, presenting the masking tone before the test tone did not disrupt pitch identification performance. Second, the perceptual processing of the test tone was interfered with by a retroactive masking tone that was presented to the ear contralateral to the test tone presentation. Third, the frequency similarity of the masking and test tones was not a critical factor in the present experiments. These results indicate that the interference tone prevents or limits the perceptual processing of the image of the test tone. The results also support the hypothesis that the auditory image is located centrally rather than at the receptor level. It remains to be seen whether the processing of a nonauditory stimulus (e.g., visual) could function as a masking stimulus in the present task.

In the present studies the results indicate that the auditory image of the test tone usually lasted about 250 msec. However, the observed time course of the image will probably be dependent upon the perceptual task of *S*. If the amount of information in an image decreases as the image decays, the duration of the image will appear to decrease as the difficulty of the task is increased. That is, the features necessary for a difficult identification will not be available in a degraded image. On the other hand, simple identifications, such as stating whether or not a tone occurred, could be possible with a degraded image.

At first glance, one study seems to give evidence about the time course of the image for easy tasks. Erikson and Johnson (1964) showed that *S*s could sometimes report the occurrence of a tone while they were reading even if questioned 10 sec. after the tone presentation. In that study, however, *S*s could have identified the tone when it was presented. Accordingly, the recall decision could have been dependent on memory for the covert identification rather than memory for the physical image of the tone. This conjecture agrees with the result that increasing the level of background noise during the retroactive interval did not interfere with performance. Thus, no direct evidence is presently available on the relationship between the time course of the image and the difficulty of the perceptual task.

Two other studies of the retroactive masking of perceptual processing of easily detectable tones are relevant to the present experiments. First, an experiment of Gol'dburt (1961) has shown that a retroactive tone shortened the perceived duration of a test tone relative to the test tone being presented alone. The masking of the perceived duration was especially effective at short durations (5–200 msec.)

of the test tone. Furthermore, with a silent interval of 240 msec. between the two tones, very little, if any, masking of the perceived duration was observed. These results were interpreted as evidence of auditory sensation continuing after termination of the stimulus producing it (Gol'dburt, 1961). The present results support this interpretation and provide further evidence about the location, time course, and stability of the auditory image (sensation) that remains after the stimulus producing it.

Second, an experiment of nonverbal or perceptual memory is relevant to the present study. Massaro (1970) has shown that a silent interval after a tone to be remembered increased the memory strength of that tone when the tone's duration was brief (200 msec.). However, at a longer duration of the test tone (500 msec.), a silent interval after the tone had no significant effect on memory performance. These results indicate that the perceptual processing during the silent interval can improve memory storage of the perceptual stimulus as well as enhance its identification.

In conclusion, the present experiments provide evidence for an auditory image that retains the sensory information of a stimulus for a short period of time after the stimulus is terminated. Furthermore, this image can be processed, while it lasts, in order to identify the stimulus that produced the image.

REFERENCES

Averbach, E., & Coriell, A. S. Short-term memory in vision. *Bell System Technical Journal*, 1961, **40**, 309–328.

Erikson, C. W., & Johnson, H. J. Storage and decay characteristics of nonattended auditory stimuli. *Journal of Experimental Psychology*, 1964, **68**, 28–36.

Fletcher, H., & Munson, W. A. Loudness, its definition, measurement, and calculation. *Journal of the Acoustical Society of America*, 1933, **5**, 82–108.

Geldard, F. A. *The human senses.* New York: Wiley, 1953.

Gol'dburt, S. N. Investigation of the stability of auditory processes in micro-intervals of time (New findings on back masking). *Biophysics*, 1961, **6**, 809–817.

Green, D. M., & Swets, J. A. Signal detection theory and psychophysics. New York: Wiley, 1966.

Homick, J. L., Elfner, L. F., & Bothe, G. G. Auditory temporal masking and the perception of order. *Journal of the Acoustical Society of America*, 1969, **45**, 712–718.

Kahneman, D. Method, findings and theory in studies of visual masking. *Psychological Bulletin*, 1968, **70**, 404–425.

Massaro, D. W. Consolidation and interference in the perceptual memory system. *Perception and Psychophysics*, 1970, **7**, 153–156.

Massaro, D. W. Perceptual processes and forgetting in memory tasks. *Psychological Review*, in press.

Neisser, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

Netter, F. H. *The Ciba collection of medical illustrations.* Vol. 1. *Nervous system.* Summit, N. J.: Ciba, 1967.

Norman, D. A. Toward a theory of memory and attention. *Psychological Review*, 1968, **75**, 522–536.

Posner, M. I., & Keele, S. W. Decay of visual information from a single letter. *Science*, 1967, **158**, 137–139.

Sperling, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, **74**(11, Whole No. 498).

Weisstein, N., & Haber, R. N. A U-shaped backward masking function in vision. *Psychonomic Science*, 1965, **2**, 75–76.

Wickelgren, W. A., & Norman, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, **3**, 316–347.

(Received April 2, 1970)