Journal of Experimental Psychology:
Human Perception and Performance
1979, Vol. 5, No. 1, 110-128

Similarity Effects in Backward Recognition Masking

Howard J. Kallman and Dominic W. Massaro
University of Wisconsin

Auditory backward recognition masking refers to the ability of a masking sound to terminate further perceptual resolution of a test sound presented slightly earlier in time. The present experiments were conducted to determine whether mask/test tone similarity effects in backward recognition masking could be reliably demonstrated. Although similarity effects were found in Experiments 1 and 2, only about 60% of the subjects demonstrated these effects. Experiment 3 was designed to isolate which stage of information processing is responsible for similarity effects. It was hypothesized that similarity effects are due to mask interference with the synthesized auditory memory of the test tone rather than to selective overwriting of a preperceptual auditory store; previous research has shown that interference in synthesized auditory memory depends on the similarity of the interfering stimulus to the items held in memory. By independently varying the backward masking interval and the interfering effect of the mask on the test tone memory, it was possible to demonstrate that similarity effects are indeed caused by mask interference in synthesized memory. The implications of these results are considered in the framework of auditory and visual masking.

When two sounds are presented to a listener, one shortly after the other, perceptual resolution of the first sound depends, in part, on the stimulus onset asynchrony (SOA) separating the two sounds. For example, it has been shown that identification of a target sound's pitch, timbre, lateralization, duration, microstructure, and loudness improves as a function of the SOA separating the target from the second sound (Hawkins & Presson, 1977; Idson & Massaro, 1977; Massaro, 1970b, 1972, 1975a, 1975b, 1976; Massaro, Cohen, & Idson, 1976; Massaro & Idson, 1976, 1977; Moore & Massaro, 1973; Pollack, 1976; Sparks, 1976); performance generally reaches asymptote at an SOA of approximately 250 msec. Because presence of the second sound degrades recognition of the first, this effect has been termed *backward recognition masking*. Massaro (1972, 1975a, 1975b, 1975c, 1976) has explained the backward recognition masking effect by assuming that a preperceptual image of a presented sound is maintained by the listener for approximately 250 msec. During this time featural information is read out at a constant rate from the preperceptual auditory store into a synthesized auditory memory. The presentation of a second sound (referred to as the *mask*) serves to terminate perceptual processing of the earlier sound. The mask interrupts the readout from the preperceptual store of information pertaining to the first sound by replacing the first sound's representation in the preperceptual store. The fact that backward recognition masking occurs regardless of whether the mask is presented ipsilateral or contralateral to the test sound suggests

This research was supported in part by National Institute of Mental Health Grant MH19399.

The authors would like to thank Gregg Oden, Willard Thurlow, Wendy Idson, and Janet Kelly for helpful discussions, and Michael Cohen and Jim Bryant for computer programming assistance. Also, thanks go to the reviewers for their constructive comments.

Requests for reprints should be sent to either Howard J. Kallman or Dominic W. Massaro, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

that the preperceptual auditory store is centrally located (Hawkins & Presson, 1977; Massaro, 1975a; Massaro & Cohen, 1975; Sparks, 1976).

Recently, certain workers have claimed that the similarity relationship between the test sound and the masking stimulus determines the magnitude of the effect of the backward mask (Holding, Loeb, & Yoder, 1972; Loeb & Holding, 1975; Sparks, 1976). These researchers contend that the magnitude of the backward mask's effect increases as the similarity of the mask to the test sound increases. The effectiveness of masking as a function of the similarity relationship between the test and masking sounds can illuminate the properties of preperceptual auditory storage and the nature of recognition masking. For example, in visual masking, the extent of backward recognition masking seems to depend, in part, on the similarity of the mask to the target stimulus (Coltheart & Arthur, 1972; Hellige, Walsh, Lawrence, & Cox, 1977; Turvey, 1973). These results may mean that the mask selectively overwrites portions of the preperceptual visual store rather than overwriting the information in the preperceptual store in toto.

Although Massaro has argued that in auditory backward recognition masking, the mask serves to overwrite *all* information held in the preperceptual store, the findings of similarity effects in auditory backward recognition masking suggest that analogous to the interpretation of the visual backward recognition masking results, the mask may only selectively overwrite information held in a preperceptual auditory store. However, many of the studies that have demonstrated mask/test tone similarity effects have not adequately addressed the question of whether these effects are due to operations at the initial perceptual stage of processing or whether they depend instead on interference effects in tone memory.

Although Massaro has argued that the backward mask serves to terminate perceptual processing of an earlier presented sound, in certain experimental paradigms the mask may also interfere with memory representations crucial to performance on the experimental task. For example, Loeb and Holding (1975) instructed subjects to perform an absolute identification task. On each trial, subjects identified a test tone as high or low in pitch; the test tone was followed by either a noise or tone mask. Performance was better under the noise mask condition. Of particular importance is the fact that mask type was a blocked variable. As Massaro (1975a, 1975b) and Hawkins and Presson (1977) have argued, blocking mask type may systematically affect the memories for the two possible test tones. In order to perform the absolute judgment task in Loeb and Holding's experiment, subjects had to remember what the two possible test tones sounded like in order to be able to classify the test tone presented on a given trial as either *high* or *low*. However, the strength of the memories for these test tones might be expected to vary depending on the amount and type of interference presented during each block of trials; less interference might be expected when noise rather than a tone was used as the mask. Thus, memory of the two test tones may have been relatively degraded on tone mask blocks; if this were the case, identification performance on those blocks should have suffered. Similarly, the experimental paradigms used by Sparks (1976) and Holding et al. (1972) do not rule out the possibility that the similarity effects they found were due to memory interference rather than to selective overwriting of a preperceptual store (see Kallman & Massaro, Note 1, for a more complete discussion). This possibility is particularly appealing given that there is evidence that interference in tone memory depends on the similarity of the interfering stimulus to the test items held in memory (Deutsch, 1974; Massaro, 1970c).

The first experiment of the present investigation was designed to minimize the degree to which any mask/test tone similarity effects found could be attributed to memory factors. The subject's task was to determine whether a test tone was higher or lower in frequency than a 1,000-Hz standard, which on each trial preceded the test tone by 500 msec. The test tone was 1,000 ± Δf Hz and was followed after a variable silent interval by one of seven masking tones ranging from 400 to 2,510 Hz. There were also trials on which no mask was presented. Rather than blocking conditions, experimental conditions occurred randomly within each session. Both the use of a fixed-frequency standard on each trial and the fact that mask frequency was varied within experimental blocks minimized the degree to which memory factors could have contributed to similarity effects.

A secondary purpose of the first experiment was to provide data relevant to what Hawkins and his colleagues (Hawkins & Presson, 1977; Hawkins, Thomas, Presson, Cozic, & Brookmire, 1974) have called the "error imbalance effect." Error imbalance refers to the finding that categorization of the test tone in backward recognition masking experiments is sometimes biased by the frequency of the masking tone. For example, if a subject is asked to judge the pitch of a target tone, his/her response may be biased toward categorizing the target as *high* given a masking tone of relatively high frequency (pitch). To explain this finding, Hawkins and Presson (1977) suggested that the masking tone integrates into the test tone percept according to an additive rule. More specifically, it was suggested that categorization of the test tone in a backward masking experiment reflects processing of all the auditory events occurring within a more or less fixed sampling interval. For example, suppose that a 20-msec test tone with frequency of 1,000 Hz was followed after a 200-msec silent interval by a 20-msec mask of frequency 1,100 Hz. Given a sampling interval of 300 msec, the categorization of the test tone would reflect 220 msec of processing of the 1,000-Hz test tone (processing of the test tone would begin with its presentation and would terminate with mask presentation) plus 80 msec of processing of the 1,100-Hz mask. Presence of the mask in this case would bias the categorization of the test tone upward (that is, toward *high*). Additionally, as the silent interval between the test tone and mask becomes smaller, the extent of the bias should become greater, for in this case a greater portion of the sampling interval would be devoted to processing the mask. Furthermore, Hawkins et al. have suggested that if such an additive process underlies the error imbalance effect, the effect should increase not only as the silent interval between test tone and mask decreases but also as the frequency separation between the test tone and mask increases. To a limited extent, the results of Hawkins et al. support this prediction. However, these studies examined a relatively small range of mask frequencies, that is, from approximately 750 to 900 Hz. Because Experiment 1 used masks ranging from 400 to 2,510 Hz, a more rigorous test of Hawkins et al.'s predictions could be made.

Experiment 1

Method

Subjects. Nineteen students participated for 5 consecutive days in order to obtain extra points in an introductory psychology course.[1] Subjects were tested in groups of four or fewer and were seated in individual sound-attenuated rooms.

Stimuli and procedure. Each trial of the experiment consisted of a standard 20-msec 1,000-Hz tone followed by a 20-msec test tone. The silent interval between the standard and test tone was 500 msec. Except on no-mask trials, a 100-msec, variable-frequency masking tone followed the test tone by a variable interval. The subject's task was to identify (by pushing one of two buttons) the test tone as higher or lower in pitch than the standard tone. The response period lasted 2 sec and began at the offset of the test tone. Feedback was presented after the response period by illuminating for 250 msec the symbol H or L, depending on whether the pitch of the test tone was higher or lower than that of the standard. Following feedback, there was an intertrial interval of 1.5 sec.

The frequencies of the test tones were adjusted between sessions to keep overall performance of as many subjects as possible at approximately 75% correct. On each trial, the test tone was either 1,000 + Δf Hz or 1000 − Δf Hz. Masking tones were 400, 630, 1,000 − 2Δf, 1,000, 1,000 + 2Δf, 1,590, or 2,510 Hz. On mask trials, onset of the masking

[1] Due to a mechanical failure, data from one subject were not recorded on one day of the experiment. This subject participated in a make-up session on Day 6.

tone followed offset of the test tone by an interstimulus interval (ISI) of 10, 20, 40, 80, 160, 250, 350, or 500 msec. On one ninth of the trials, no masking tone was delivered (the frequency of the masking tone was a dummy variable under the no-mask condition; the no-mask condition was a level of ISI). The 126 experimental conditions ($2 \times 7 \times 9$ [test tone] $\times$ [masking tone] $\times$ [ISI]) occurred with equal probability according to a random schedule without replacement.

On Day 1, 10 blocks of 50 trials each were administered. Initially, Δf was set to 40 Hz; Δf was subsequently adjusted between blocks throughout the experiment to achieve an average performance of approximately 75% correct. Since changes in Δf were made between blocks of trials, all experimental conditions had equivalent Δfs. On Day 2, 5 blocks of 50 trials were followed after a short break by a single block of 262 trials. On Days 3 and 4, 2 blocks of 262 trials were presented, and on Day 5, 3 blocks of 262 trials were presented. For blocks of 262 trials, the first 10 trials were not scored, and sampling without replacement began anew with the 11th trial. Only data from Days 3, 4, and 5 were included in data analysis. Thus, for each subject, there were 14 observations at each experimental condition.

All tones in the experiment were sine waves generated by a digitally controlled oscillator (Wavetek Model 155). Tones were delivered binaurally to subjects at 80 dB(A) through Grason-Stadler TDH-49 headphones. The tones began at the zero crossing and reached maximal intensity in 1/4 of a cycle. Feedback was presented over a visual display of light-emitting diodes (Monsanto MDA-III). Experimental events and data collection were controlled by a PDP-8/L computer.

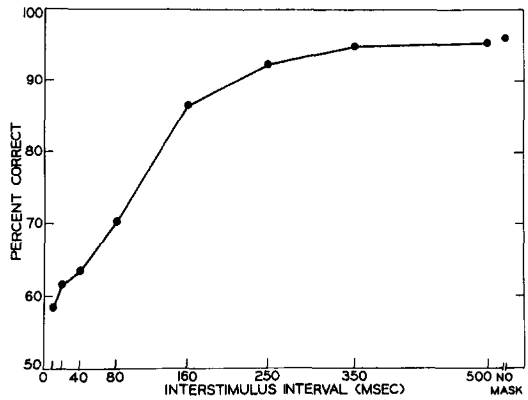

Figure 1. The percentage of correct identifications of the test tones as a function of the duration of the silent interstimulus interval—Experiment 1.

Results and Discussion

Only data from subjects whose overall performance was between 65% and 90% correct on 1,000-Hz mask trials were included in data analysis. Data from seven subjects whose performance was below 65% correct were excluded; no subject achieved greater than 90% correct. In addition, one subject demonstrated a substantial bias to respond "High" regardless of experimental condition and his data were eliminated; this subject responded "High" on 62% of the trials. Statistics were thus performed on a total of 11 subjects.

An initial analysis of variance (ANOVA) was performed to evaluate the effect of ISI. ISI, test tone (i.e., high or low), block, and mask frequency were within-subject variables. As can be seen in Figure 1, performance improved from approximately 60% to 95% correct with increases in ISI; this main effect of ISI was highly significant, $F(8, 80) = 105.55$, $p < .001$. Visual inspection of data from individual subjects indicated that every subject demonstrated a backward recognition masking function approximately comparable to that shown in Figure 1. Because mask frequency was a dummy variable under the no-mask level of ISI, a subsequent ANOVA was performed in which the no-mask level of ISI was excluded. Of particular interest was the effect of mask frequency. Figure 2 plots performance as a function of mask frequency; ISI is the curve parameter. The significant quadratic trend, $F(1, 60) = 29.14$, $p < .001$, indicates that performance decreased as a function of mask/test tone similarity. Thus, mask/test tone similarity effects were convincingly demonstrated in Experiment 1. However, mask/test tone similarity effects occurred only at the shorter ISIs, as is confirmed by a significant interaction between ISI and mask frequency, $F(42, 420) = 3.50$, $p < .001$. However, the absence of similarity effects at the longer ISIs may have been due to the high overall performance levels achieved at these ISIs. Frequency per se did not determine the effectiveness of the masks, as confirmed by the absence of a significant linear trend across mask frequencies, $F(1, 60) = 2.31$, $p > .10$. The residual variance after partitioning the linear

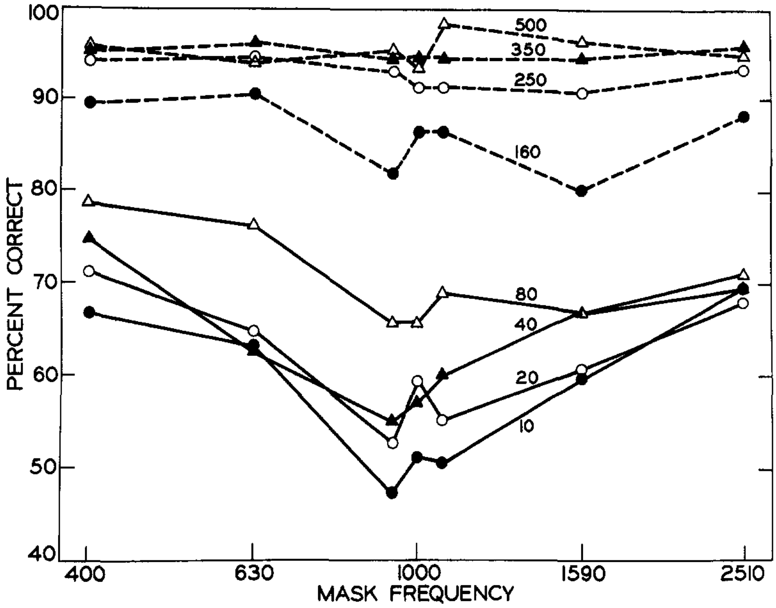

Figure 2. The percentage of correct identifications of the test tones in Experiment 1 as a function of mask frequency. (ISI, in msec, is the curve parameter. Note that mask frequency is plotted on a log scale.)

and quadratic mask components was significant, $F(4, 60) = 5.28$, $p < .001$; however, given that the linear and quadratic trend components captured most of the variance, tests for higher order trends were not pursued.

The percentages of correct data for each subject at each mask frequency are presented in Table 1; these data are summed over ISIs. As is evident from examination of this table, notable individual differences were found. Each subject's data either showed pro-

Table 1
Percentage Correct at Each Mask Frequency (Summed Over ISIs of 10–500 msec) for Individual Subjects in Experiment 1 and Results of Tests for Quadratic Trends, with the Mean Δfs Used When Presenting Test Tones to Subjects

Subject	Mask frequency							F_{quad}	MS_e	Δf
	400	630	1,000 − 2Δf	1,000	1,000 + 2Δf	1,590	2,510			
1	97	88	75	82	75	92	95	116.40**	.036	39
5	91	87	73	68	75	83	88	43.51**	.079	70
6	90	84	78	80	79	86	91	28.51**	.060	70
7	93	88	76	75	76	88	91	52.27**	.056	64
8	89	83	74	76	74	80	89	45.06**	.059	44
2	75	76	69	69	73	69	75	2.10	.109	39
3	71	72	69	67	71	65	61	1.52	.071	70
4	79	81	78	80	75	74	82	4.33*	.032	70
9	81	72	75	79	80	68	78	1.72	.063	44
10	69	75	68	75	75	69	68	4.60***	.049	96
11	79	76	77	74	79	72	76	$F < 1$	.073	96
M	83	80	74	75	76	77	81			64

Note. Degrees of freedom for F_{quad} were 1, 36.
* $p < .05$.
** $p < .001$.
*** $p < .05$, but trend is negative.

nounced mask/test tone similarity effects or an absence of an effect of similarity. Single-subject ANOVAs were performed; replications (blocks) served as the random factor. The results of tests for quadratic trends across the mask frequencies appear in Table 1.[2] Nearly all the subjects either demonstrated a highly significant quadratic trend ($n = 5$) or a clear absence of such a trend ($n = 5$). Thus, the statistical analyses confirm the individual differences. Also presented in Table 1 are the mean Δfs used when presenting the test tones to subjects. Inspection of Table 1 makes clear that whether a subject demonstrated similarity effects did not critically depend on the size of Δf.

Figure 3 shows the interaction between test tone and mask frequency. Error imbalance was clearly found as the significant interaction confirms, $F(6, 60) = 3.84$, $p < .01$. Error imbalance decreased with increases in ISI, $F(42, 420) = 2.86$, $p < .001$, but ceiling effects might have obscured the error imbalance effects at the longer ISIs. Figure 3 also indicates that Hawkins et al.'s (1974) and Hawkins and Presson's (1977) prediction that the error imbalance effect should increase with greater frequency separations between masks and test tones was not borne out. In fact, the error imbalance effect was greater for relatively near as opposed to far masks. This finding calls into question Hawkins et al.'s additive model of backward masking. It might be possible to salvage the additive model by assuming that the error imbalance effect occurs maximally when test tone and mask occur within the same channel (defined perhaps by frequency range); this would explain why error imbalance was greater with near as opposed to far masks. Such a qualification would still enable one to assume that within a channel, the magnitude of error imbalance would depend on an additive process. However, the results of Experiment 3 below suggest that Hawkins et al.'s additive model needs still further revision.

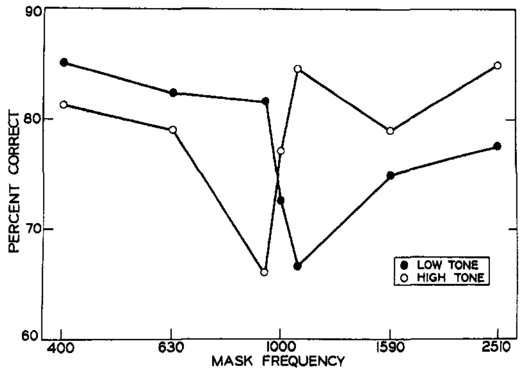

Figure 3. The percentage of correct identifications of the low and high test tones as a function of mask frequency in Experiment 1. (The data are summed over ISIs of 10–500 msec.)

Experiment 2

Mask/test tone similarity effects were demonstrated in Experiment 1 when mask/test tone similarity was varied along the frequency dimension. Experiment 2 was an attempt to extend the finding of similarity effects to the case in which a test tone is followed by either another tone or instead by band-limited noise. The noise and tone masks used in Experiment 2 might be considered more fundamentally different from one another than the different frequency tone masks used in Experiment 1. It was thus of interest to determine whether the individual differences found in Experiment 1 would also appear in Experiment 2. Given the more extreme manipulation of similarity, it seemed possible that all subjects might show similarity effects.

Method

Subjects. Ten students participated for 4 consecutive days in order to obtain extra points in an introductory psychology course. Subjects were tested in groups of four or fewer and were seated in individual sound-attentuated rooms.

Stimuli and procedure. On the first day of the experiment, subjects were initially presented with a block of 298 single-tone, absolute-identification trials. The purpose of these trials was to familiarize subjects with the high and low tones. The task was simply to indicate by pressing a button whether the tone was low or high in pitch. The low and

---

[2] In computing the sums of squares for trends, frequency was considered on a log scale. For mask frequencies of 1,000 ± 2Δf, the grand mean Δf (over subjects) was used for computing the trend coefficients used on all tests.

high tones had frequencies of 950 and 1,040 Hz, respectively, and these occurred randomly without replacement. The single-tone, absolute-identification trials were essentially identical to the no-mask experimental trials described below.

An experimental trial consisted of a high- or low-frequency 20-msec test tone followed (except on no-mask trials) after a varying ISI by a 100-msec sound, which was either a 1,000-Hz tone or narrow-band noise. Test tones were $1{,}000 \pm \Delta f/2$ Hz,[3] and the subject's task was to identify, by pressing a button, the tone as either high or low in pitch. The response period lasted 2 sec and began after offset of the test tone. Following the response period, visual feedback indicating whether the test tone had been high or low was given. Intertrial interval after feedback was 1.5 sec.

Following presentation of the single-tone, absolute identification trials on Day 1, five blocks of 50 experimental trials were administered. Initially, Δf was set to 120 Hz, but this was varied between blocks throughout the experiment to keep subjects' overall performances at approximately 75% correct. On Days 2, 3, and 4, two experimental blocks of 298 trials were presented; each experimental block was preceded by 50 single-tone, absolute-identification trials. The first 10 of the 298 experimental trials were not recorded for data analysis; sampling began anew with the 11th trial. The Δf on single-tone trials was set to the Δf used for the subsequent experimental block.

ISI from the offset of the test tone to the onset of the mask was 10, 20, 40, 80, 160, 250, 350, or 500 msec. On one ninth of the trials, no masking tone was presented (mask type was a dummy variable under the no-mask level of ISI). The 36 experimental conditions ($2 \times 2 \times 9$ [test tone] $\times$ [mask] $\times$ [ISI]) occurred randomly without replacement. For data analysis, there were 48 observations per subject at each experimental condition.

Method of stimulus control and presentation was the same as that described for Experiment 1 except as noted below. Unlike Experiment 1, in which all sounds were sinusoids, half of the masks in Experiment 2 were narrow-band noise centered at 1,000 Hz with a 24 dB/octave falloff. Two samples of narrow-band noise were digitized for playback during the experiment. White noise generated by a Grason Stadler noise generator was both high and low pass filtered at 1,000 Hz using a Krohn-Hite model 3500R filter. The noise was read into computer memory through analog-to-digital converters with the aid of a program designed to store samples at a 10K sampling rate. During the experiment, the noise was output through digital-to-analog converters and was low and high pass filtered at 3,000 Hz and 30 Hz, respectively. The resulting narrow-band noise as presented to subjects was centered at 1,000 Hz with approximately a 24 dB/octave falloff. The tones generated by the Wavetek oscillator and the filtered noise were fed into a Grommes Precision Mixer (model ST-6) prior to amplification through a McIntosh Model MC-50 amplifier. Noise and tones were presented at 81 dB(A). Throughout the experiment, a given subject was always presented the same noise sample on noise mask trials.

Results and Discussion

Data from Day 1 were disregarded, leaving the data from Days 2, 3, and 4 for analysis. Only data from subjects whose overall percent correct on experimental trials on Days 2, 3, and 4 was between 65% and 90% were included in data analysis. A total of four subjects did not meet these criteria; of the four, two each were at the ceiling and the floor. As a result, data from six subjects are reported below.

An analysis of variance was performed on the data. ISI, mask type (i.e., noise or tone), test tone (high or low), and block were within-subject variables. Noise sample (two levels, three subjects per sample) was a between-subject variable. Performance at each ISI is plotted in Figure 4; mask type is the curve parameter. Overall, performance improved from approximately 60% to 90% correct with increases in ISI; this main effect of ISI was highly significant, $F(8, 32) = 70.90$, $p < .001$. Thus, as in Experiment 1, backward recognition masking was convincingly demonstrated. Because mask type was a dummy variable under the no-mask level of ISI, a subsequent ANOVA was performed in which the no-mask level of ISI was omitted. Although performance was better with a noise than a tone mask, this difference only approached significance, $F(1, 4) = 4.70$, $.05 < p < .10$. However, mask type interacted with ISI, $F(7, 28) = 3.34$, $p < .025$, as can be seen in Figure 4. Noise masks degraded performance less than did tone masks, but only at the shorter ISIs.

[3] In some cases, it was necessary for $\Delta f/2$ to be asymmetric around 1,000 Hz. The Wavetek oscillator can be adjusted by units of 10 Hz when the frequency is set above 1,000 Hz but by units of 1 Hz when set to a value below 1,000 Hz. Thus, if the Δf needed to keep performance at 75% correct was not a multiple of 20 Hz, the frequency of the lower test tone was adjusted accordingly.

The effect of noise sample was not significant, $F < 1$, nor was the interaction between noise and mask type, $F(1, 4) = 1.80$, $p > .25$, or the interaction between noise, mask type, and ISI, $F(7, 28) = 1.65$, $p > .10$.

Individual subject masking functions are plotted in Figure 5; the curve parameter is mask type. As in Experiment 1, there appear to be individual differences with regard to the effect of mask type, although these differences are perhaps less pronounced than in Experiment 1. Subjects 5 and 6 demonstrated pronounced mask/test tone similarity effects, whereas Subjects 2 and 3 showed clear but less dramatic effects of similarity. Subjects 1 and 4 did not show an effect of mask/test tone similarity.

The results of Experiment 2 are consistent with the major findings of Experiment 1. Despite the different mask types used between the two experiments, mask/test tone similarity effects were found in both. Furthermore, individual differences with regard to the effect of mask type were found in both experiments.

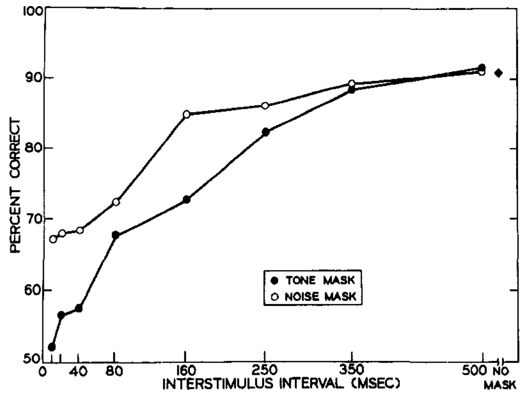

Figure 4. The percentage of correct identifications of the test tones in Experiment 2 as a function of the duration of the silent interstimulus interval. (Mask type is the curve parameter.)

Experiment 3

Given the results of Experiments 1 and 2, it is tempting to conclude that the similarity effects found were due to selective overwriting of a preperceptual auditory store rather than to memory factors. In these experiments, precautions were taken to insure that the quality of the memory representations of the sounds that the test tones were to be compared to were equivalent for each masking stimulus condition. However, one possibility not yet considered is that in addition to terminating perceptual processing of the test tone, the mask might interfere with memory for the test tone *while a decision about the test tone is being made.* Experiment 3 was designed to test the possibility that mask/test tone similarity effects are due to such interference in memory.

Before describing the details of Experiment 3, we present our conceptualization of how mask interference with the test tone memory might affect performance accuracy. According to Massaro's theory, perceptual processing of a sound continues until either another sound is presented or approximately 250 msec have elapsed. We assume, then, that given an SOA of less than approximately 250 msec, a subject would not normally arrive at a *final* decision about the test tone until after mask onset, since the readout of information from the preperceptual store would not be complete until mask onset. But although the final decision about the test tone would not be made until sometime after mask onset, a comparison process that compares the test tone to an earlier presented standard (or in an absolute identification task to memories of how the high and low test tones sound) could begin as soon as *any* information about the test tone has been read out from the preperceptual store. Thus, we assume that the test tone comparison process begins almost immediately after test tone presentation (since information from the preperceptual store begins to be read into a synthesized auditory memory at this time), but the final decision about the test tone does not occur until sometime after completion of the readout of test tone information from the preperceptual store. The final decision then depends on the sum of information derived from a comparison process that begins shortly after test tone presentation and extends until sometime after presentation of the mask.

A key point in our conceptualization is that as soon as the mask is presented, the readout from the preperceptual store of information about the mask begins, and as this information is read out, it interferes with the memory of the test tone. After mask onset, the test tone comparison process must thus rely on a test-tone-synthesized memory that has been degraded by mask interference; prior to mask onset, the comparison process relies on a test tone representation free from such interference. As a result, the degree to

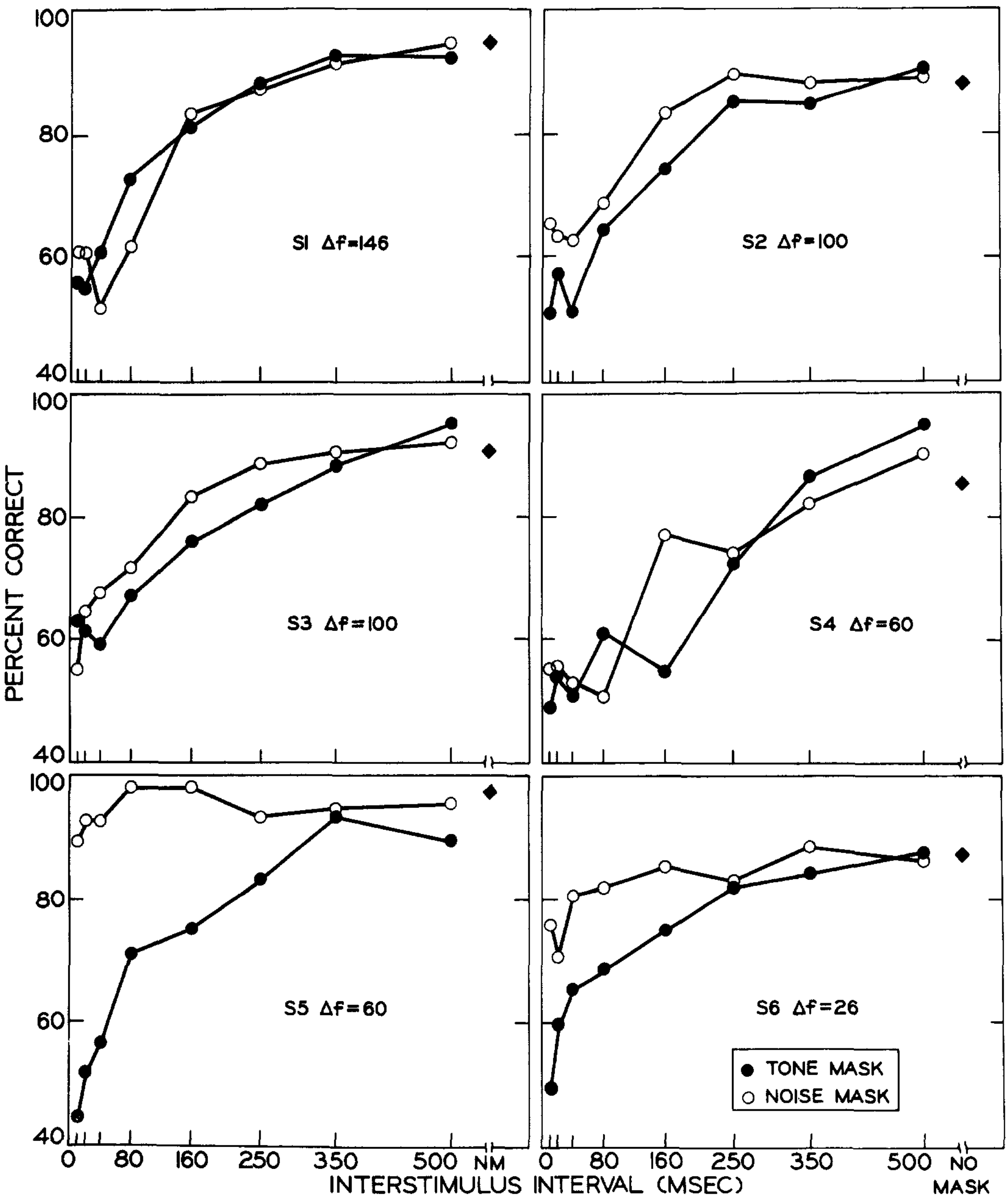

Figure 5. Individual subject data from Experiment 2 showing the percentage of correct identifications of the test tones as a function of the duration of the silent interstimulus interval. (Mask type is the curve parameter. The Δfs for each subject are averaged over experimental blocks.)

which mask interference will play a role in the final decision will depend on the relative amount of time that the comparison process must rely on an interference-free, synthesized memory of the test tone as compared to a synthesized memory of the test tone degraded by mask interference. This suggests that in our first two experiments, the effect that mask interference with the test-tone-synthesized memory would have had on performance should have been greatest at short ISIs because at the short ISIs, mask interference would have affected the test tone comparison process shortly after it had begun. Thus, in Experiments 1 and 2, the effect of mask interference on performance should have been greatest at short ISIs because at the short ISIs, mask interference would have substantially degraded the test tone comparison process. If mask/test tone similarity effects result from differential mask interference with the synthesized memory of the test tone, the magnitude of these effects should be greatest at the short ISIs. Although similarity effects did seem to decrease with increases in ISI in Experiments 1 and 2, we should note that this result is also predicted by the hypothesis that similarity effects are due to a selective overwriting of information held in the preperceptual store. This follows because the backward masking effect of a stimulus decreases with increases in ISI; consequently, any effect of similarity due to selective overwriting (backward masking) would also decrease with increases in ISI.

Experiments 1 and 2 do not allow us to choose whether similarity effects are due to selective overwriting of a preperceptual store or, instead, to differential mask interference with the test tone memory. Experiment 3 provides a critical test between these two alternatives. As noted above, the idea that the similarity effects found in Experiments 1 and 2 were due to mask interference with the test tone memory predicts a decrease in the magnitude of similarity effects with increases in ISI. However, note that in these experiments, the test tone comparison process could begin at the time of test tone presentation. In Experiment 1, the standard that the test tone was to be compared to was presented *prior* to test tone presentation, and thus subjects were able to start comparing the test tone to the standard as soon as the test tone was presented. Similarly, in Experiment 2 subjects had knowledge of what the high and low test tones sounded like at the time of test tone presentation and, consequently, could immediately begin comparing the test tone to the memories of these sounds.

But suppose that the reference stimulus to which the test tone is to be compared does not occur until sometime after both the test tone and the mask have been presented. In this case, the test tone comparison process could not start until presentation of the reference stimulus, that is, the comparison process could not begin until sometime after both the test tone and the mask have been presented. Given this type of paradigm, we would predict that if similarity effects are due to memory interference, the magnitude of mask/test tone similarity effects would remain approximately constant across the ISIs separating the mask and test tone. This is because, irrespective of ISI, the test tone comparison process would, during its entire duration, need to rely on a test tone memory degraded by interference. In contrast, if similarity effects were due to selective overwriting of a preperceptual store, we would expect them to decrease with increases in ISI.

Two trial types were employed in Experiment 3: mask second (M2) and mask third (M3); Figure 6 illustrates these two types of trials. The subject's task on M3 trials was to compare a standard and a test tone that were separated by 700 msec and to determine whether the test tone was higher or lower in pitch than the standard. A tone or noise mask followed the test tone by a variable ISI. On M2 trials, the task was also to judge whether a test tone was higher or lower in pitch than a standard that preceded it; however, the noise or tone mask followed the standard by a variable ISI.

Mask/test tone effects on M3 trials should decrease in magnitude as the ISI between the test tone and mask increases. As was

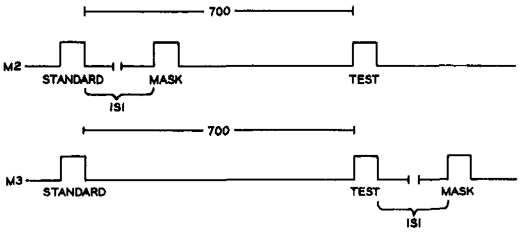

Figure 6. The sequence of events on M2 and M3 mask trials in Experiment 3. (ISI = interstimulus interval.)

the case for trials in Experiments 1 and 2, this prediction follows regardless of whether test tone memory interference or selective overwriting of a preperceptual store underlies the similarity effects. The M2 trials provide the critical test between the two hypotheses. If similarity effects are due to selective overwriting of a preperceptual store, the effect of similarity on M2 trials should decrease with increases in ISI. Again, this is because the effect of backward masking (overwriting) decreases with increases in ISI. However, if similarity effects are due to memory interference, a different prediction is made. Regardless of the ISI on M2 trials, the masking stimulus would always interfere with the synthesized memory of the standard prior to presentation of the test tone. Because interference in tone memory depends primarily on the occurrence of intervening stimuli rather than simply decay over time (Massaro, 1970a, 1970c), the amount of interference with the standard tone representation resulting from mask presentation would be expected to remain roughly constant across ISIs. Thus, during the entire time that the test tone was being compared to the standard, the comparisons between tones would rely on a memory of the standard that had been degraded by mask interference, and the amount of this interference would remain constant across ISIs on M2 trials. As a result, the memory interference hypothesis predicts that similarity effects should remain constant across ISIs. But if, on M2 trials, similarity effects decrease with increases in ISI, this would be consistent with the idea that they result from selective overwriting of the preperceptual store.

In order to assess the effect of mask type at each trial type (M2 and M3) at each ISI without the problem of different performance levels at each ISI, a computer-controlled algorithm adjusted Δf at each ISI for each trial type so that overall performance of each subject was approximately 75% correct at each ISI. Thus, Experiment 3 provided a sensitive test of the effect of similarity at each ISI for both M2 and M3 trials.

Method

Subjects. Eight students at the University of Wisconsin served for 5 consecutive days and were paid $2 per hour. One of the students had previously participated in a pilot backward recognition masking experiment, although his data from that experiment had not yet been analyzed when he was asked to participate in the present study. In addition, the first author served as a subject.

Stimuli and procedure. All sounds in the experiment were of 20 msec duration. All tones were sinusoids. The same sample of noise was presented to each subject on each noise trial.

There were two types of blocked trials—M2 and M3.[4] Both trial types began with the presentation of a standard tone followed by a test tone presented 700 msec after the offset of the standard. The subject's task was to indicate by pressing one of two buttons whether the test tone was higher or lower in pitch than the standard tone. The response period lasted 2 sec and began with offset of the test tone. Following the response period, visual feedback indicating whether the test tone was higher or lower in pitch than the standard was given for 250 msec. The following trial began 1.5 sec after the feedback terminated.

The major difference between M2 and M3 trials was that on M2 trials the mask followed the standard tone by a variable ISI, whereas on M3 trials the mask followed the test tone by a variable ISI.

On M3 trials, the frequency of the standard varied randomly from trial to trial, with the

[4] M2 and M3 trials were blocked to minimize the possibility that subjects would respond to the wrong stimuli on a given trial. Although it was argued in the introduction that in certain cases, blocking experimental conditions can result in uninterpretable data, this was not a problem in the present experiment. The primary interest was whether the form of the interaction between mask type and ISI would differ between the M2 and M3 conditions. Any differences between M2 and M3 trials that may have resulted from blocking should have affected performance at each ISI equally and thus would not have affected the form of these interactions.

restriction that its frequency was always even numbered and within the range 700–900 Hz. The frequency of the test tone was equal to the frequency of the standard on that trial plus or minus the Δf specified for the trial's ISI. The masking stimulus followed the test tone after a variable ISI and was either a 1,200-Hz tone or narrow-band noise centered at 1,200 Hz.

On M2 trials, the frequency of the test tone was randomly varied in the same way as the frequency of the standard was varied on M3 trials. That is, the frequency of the test tone on M2 trials was even numbered and was in the range 700–900 Hz. The standard tone had a frequency equal to that of the test tone on the given trial plus or minus Δf. As on M3 trials, the Δf was again defined with reference to the appropriate ISI.

It is important to emphasize that for both M2 and M3 trials, the tone that was followed by the backward mask had a frequency equal to that of the other comparison tone plus or minus Δf. Thus, the frequency parameters specifying the to-be-masked tone were calculated in the same way for M2 and M3 trials. The to-be-masked sound was the standard on M2 trials, whereas it was the test tone on M3 trials. We were interested in the mask's effect on the sound that preceded it and, consequently, we recorded and analyzed the data in a manner consistent with this. For example, on M2 trials the subject's task was to indicate whether the test tone was higher or lower in pitch than the standard. Because on M2 trials we were interested in the backward mask's effect on the standard, the subject's responses on these trials were inverted prior to data analysis. Thus, when we present data on error imbalance, a "High" response indicates that the subject thought that the standard was higher in pitch than the test tone. In fact, in this case the subject would have responded that the test tone was lower than the standard. For presentation of the data, we label the to-be-masked sound as the *target*. Thus, on M2 trials the standard acted as target, whereas on M3 trials the test tone acted as target.

On the first day of the experiment, subjects first participated in a short introductory session consisting of 64 trials. The trials were as described for the M3 blocks with the exception that only no-mask trials were presented. The Δf was equally often 10, 20, 30, 40, 50, 60, 70, or 80 Hz according to a random schedule. Following the introductory session, the experiment proper began (although data from Day 1 were disregarded in subsequent data analyses).

On a given day in the experiment, subjects participated in one M2 and one M3 block of trials. On Day 1, five subjects first participated under the M3 condition; the other subjects first participated under the M2 condition. A short break separated presentation of the two experimental blocks. For a given subject, the type of block presented first on each day was alternated from day to day.

A block of experimental trials consisted of either 288 M3 or 288 M2 trials. Within-block experimental variables for both types of blocks were: mask type (noise or tone), target (relatively high or low), and ISI (10, 20, 40, 80, 160, 250, or 350 msec). Including the no-mask condition, there were eight levels of ISI. There were thus a total of 32 within-block experimental conditions. Each block of 288 trials was divided into nine subblocks of 32 trials. Each of the 32 within-block experimental conditions occurred randomly without replacement within each subblock. Percentages of correct data from the first subblock of each block of trials were not recorded for data analysis but were used to adjust Δfs. For each trial type (M2 and M3), a total of 32 observations per subject at each experimental condition were recorded for data analysis.

Initially, on Day 1, Δf was set to 80 Hz for all ISIs. An algorithm was developed and used throughout the experiment to keep the overall performance of each subject at each ISI within each block type at approximately 75% correct.[5]

[5] The computer subroutine adjusted Δf at each ISI every 32 trials, that is, after each subblock. Following completion of a subblock, the number of correct responses at each ISI was examined. If at a given ISI three responses (75%) were found correct, Δf was left unchanged for the next subblock. If four responses were found correct, the program jumped to a routine designed to decrease the appropriate Δf. The magnitude of this decrease was determined by whether performance at the same ISI during the subblock prior to the one under consideration had been greater than 75%, equal to 75%, or less than 75% correct. Based on this information, the routine computed a number that was then subtracted from the current Δf to arrive at a new Δf. If more than 75% of the responses during the previous subblock were correct, the number to be subtracted from the current Δf was found by shifting the current Δf (which was represented in computer memory as a binary number) right two times. For example, if the current Δf was 25_{10} ($11{,}001_2$), 6_{10} (110_2) would be subtracted to arrive at a new Δf of 19_{10}. If during the previous subblock 75% of the responses were correct, the number subtracted from the current Δf was equal to the current Δf shifted right three times. And if less than 75% were correct on the previous subblock, the number to be subtracted was equal to Δf shifted right four times. Analogously, if performance at a given ISI on the trials just completed was less than 75% correct, the computer program increased the Δf by adding to the current Δf a number equal to the current Δf shifted right two, three, or four times.

If in computing the new Δf, the result of the arithmetic shift was zero, the number 1 was either added to or subtracted from (depending on whether Δf was to be increased or decreased) the current

The 800-Hz tones and the 1,200-Hz tones and noise were presented to subjects at 81 dB(A). All other tones were presented at the peak-to-peak voltage used for presenting the 800-Hz tones. Rather than filtering the white noise before feeding it into the computer memory, as was done in Experiment 2, the white noise was fed directly into the computer memory via the analog-to-digital converter. When played back to subjects, the noise was high and low pass filtered at 800 Hz using the Krohn-Hite Model 3500R filter. The band-limited noise was thus centered at 800 Hz with an approximately 24 dB/octave falloff. Subjects were tested individually in sound-attenuated rooms.

Results and Discussion

There were two dependent variables in Experiment 3: Δf and percentage correct. Log mean Δfs for individual subjects at each ISI for each of the two mask position conditions appear in Table 2. The log mean Δfs were subjected to an ANOVA. Within-subject factors were ISI (eight levels including the no-mask condition), and mask position (M2 and M3). As can be seen in Table 2, Δf generally decreased with increases in ISI for both M2 and M3 trials. This main effect of ISI was significant, $F(7, 56) = 15.17$, $p < .001$, and indicates that backward recognition masking was evident.[6] Mask position did not interact with ISI, $F(7, 56) = 1.41$, $p > .10$; backward recognition masking was demonstrated under both the M2 and M3 conditions.

Log mean Δfs for M2 and M3 trials were 44 and 27 Hz, respectively, $F(1, 8) = 15.23$, $p < .01$. The major difference between the Δfs obtained for M2 and M3 trials is that at each ISI, the Δf was greater on M2 trials; however, Δf decreased with increases in ISI for both M2 and M3 trials. The greater Δfs on M2 trials confirms that the mask interfered with the synthesized memory of the standard tone. The fact that on M2 trials Δf dropped from 41 Hz at an ISI of 350 msec to 21 Hz under the no-mask condition further confirms the interfering effect of the mask on these trials. In contrast, the decreases in Δf with increases in ISI on both M2 and M3 trials are evidence for backward masking. In general, the results of the Δf analysis support the distinction between backward masking and synthesized memory interference.

The percentages of correct data from Experiment 3 were subjected to ANOVAS. An initial ANOVA was performed to determine whether overall performance remained constant across ISIs. Within-subject variables were: day (four levels), ISI (eight levels), mask type (two levels), mask position (two levels), and target (for M3 trials, the target variable refers to whether the test tone presented on a trial was higher or lower in frequency than the standard; on M2 trials, the target variable refers to whether the standard tone was higher or lower in frequency than the test tone). Figure 7 shows that performance at each ISI was approximately 75% correct for both M2 and M3 trials, although the effect of ISI was marginally significant, $F(7, 56) = 2.31$, $p < .05$.

Since mask type was a dummy variable under the no-mask level of ISI, the no-mask condition was disregarded in subsequent analyses. Overall, subjects correctly identified the test tone on 72% of the tone mask

---

Δf to arrive at the new Δf. The minimum and maximum Δfs allowable were 1 and 350 Hz, respectively. In the event that these values were reached, the program kept Δf at the limiting value until performance dictated a change toward a more intermediate Δf value.

[6] It is a bit puzzling that two subjects (Subject 2 and Subject 6) who achieved very small Δfs did not demonstrate backward masking in Experiment 3. Thus, these subjects were asked to participate in a subsequent backward recognition masking experiment. In this experiment, a trial began with presentation of a 20-msec, 800-Hz standard. The standard was followed after a 700-msec silent interval by a 20-msec test tone with a frequency of 800 Hz plus or minus 4, 6, or 9 Hz. A 1,200-Hz, 20-msec tone mask followed the test tone by a variable interval. The subject's task was to indicate whether the test tone was higher or lower in pitch than the standard. The percentages of correct data from both subjects showed evidence of backward masking. For example, with a Δf of 6 Hz, Subject 2's performance ranged from 61% correct at an ISI of 10 msec to 72% correct at an ISI of 350 msec. At the same ISIs, Subject 6 achieved 75% and 97% correct, respectively. It is unclear why these subjects failed to demonstrate backward masking in Experiment 3.

Table 2: *Log Mean Δfs (Hz) at Each Interstimulus Interval (ISI) in Experiment 3*

		ISI (in msec)							
Subject	Group[a]	10	20	40	80	160	250	350	No mask
					Mask-third trials				
1	N	127	161	114	75	75	68	63	50
2	N	9.1	8.8	7.5	6.4	8.4	7.7	4.9	5.6
3	S	48	53	42	26	18	21	21	13
4	S	38	38	36	31	25	27	31	30
5	S	46	29	29	13	12	11	10	7
6	N	6.2	7.2	5.3	5.9	6.0	5.7	5.4	4.7
7	S	79	79	74	56	40	35	49	28
8	S	53	83	71	80	48	60	45	27
9	S	95	78	54	52	51	55	59	39
log *M*		39	41	34	27	23	24	22	17
					Mask-second trials				
1	N	206	202	240	165	209	212	174	71
2	N	6.7	7.9	6.2	8.2	7.6	6.4	7.9	7.0
3	S	90	93	83	37	27	31	32	18
4	S	67	58	58	73	40	53	40	50
5	S	63	43	20	17	16	10	9	8
6	N	9.1	12.5	11	8.6	7.0	9.8	10.1	2.9
7	S	70	105	128	74	66	66	67	19
8	S	267	286	276	130	199	222	276	54
9	S	128	129	141	113	120	90	110	84
log *M*		61	64	59	44	41	40	41	21

[a] Subjects in Group S demonstrated mask/test tone similarity effects, subjects in Group N did not.

trials and on 79% of the noise mask trials, $F(1, 8) = 8.56$, $p < .025$. Because the main purpose of Experiment 3 was to compare the effect of mask/test tone similarity under the M2 and M3 conditions, subjects were divided into two groups according to whether their data demonstrated similarity effects. Each of the six subjects assigned to Group S were assigned on the basis of having demonstrated at least a 4% advantage (averaged over ISIs of 10–350 msec) for noise as compared to tone mask trials under both mask position conditions. Each of the three subjects assigned to Group N did not demonstrate greater than a 1% advantage for noise masks under either mask position condition. Separate analyses were performed on Group N and Group S data; the results of the Group S analyses are presented first.

Figure 8 illustrates the effect of mask type for Group S subjects as a function of ISI for M3 and M2 trials. It is important to keep in mind that the Δfs were adjusted to keep overall performance summed over mask type at each ISI at 75% correct. As expected, performance was better with noise rather than tone masks, $F(1, 5) = 32.85$, $p < .01$. However, the effect of mask type generally decreased with longer ISIs, $F(6, 30) = 2.75$, $p < .05$. Most important, though, was the significant interaction between mask position, mask type, and ISI, $F(6, 30) = 3.20$, $p < .025$. On M3 trials, similarity effects decreased with increases in ISI, but on M2 trials, the effect of mask/test tone similarity remained approximately constant across ISIs. Separate ANOVAS conducted on the data from M3 and M2 trials, revealed a significant interaction between mask type and ISI for M3 trials, $F(6, 30) = 3.94$, $p < .01$, but not for M2 trials, $F(6, 30) = 1.14$, $p > .25$. The fact that on M2 trials the similarity effects remained constant across ISIs argues that these effects are not due to selective overwriting of a preperceptual store. Instead, the results indicate that

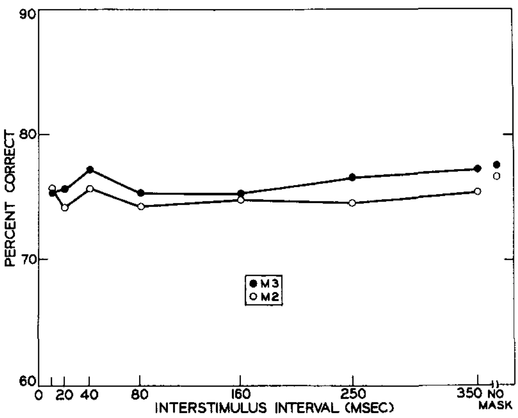

Figure 7. The percentage of correct identifications of the target tones in Experiment 3 as a function of the silent interstimulus interval. [Trial type (M2 or M3) is the curve parameter.]

mask interference with test tone memory is responsible for the similarity effects.

Presented in Figure 9 are Group S data showing the interaction between target and mask type, $F(1, 5) = 13.56$, $p < .025$. This two-way interaction was complicated by both a three-way interaction between target, mask type, and ISI, $F(6, 30) = 3.69$, $p < .01$, and a four-way interaction between target, mask type, ISI, and mask position, $F(6, 30) = 6.49$, $p < .001$. These interactions indicate the presence of an error imbalance effect of the sort noted by Hawkins et al. (1974) and Hawkins and Presson (1977). Performance was relatively poor when a low target was followed by the 1,200-Hz tone mask. Apparently, presentation of the 1,200-Hz tone mask resulted in a bias to identify the target as *high*. The bias remained constant across ISIs on M2 trials but diminished appreciably across ISIs on M3 trials. In fact, at the longer ISIs on M3 trials, a bias to identify the target as *low* on tone mask trials was evident.

The interaction between target, ISI, and mask position on the tone mask trials is important, for it addresses itself to the mechanisms underlying the error imbalance effect. If an additive model based on operations at the perceptual stage of processing were the basis for the error imbalance effect, error imbalance would be expected to decrease with increases in ISI on both M2 and M3 trials. However, if error imbalance were based on operations after the perceptual stage of processing, an interaction of the form found here would be expected. For example, rather than assuming that a sound and the mask that follows it integrate to form a percept, it might be assumed that the synthesized memory of the sound preceding the mask is biased by presentation of the mask. If this were the case, it would not matter on M2 trials whether the bias was introduced 10 msec or 200 msec after presentation of the first sound, for in either case the synthesized auditory memory representation of the first sound would become biased prior to the time that the sound that it is to be compared to (i.e., the test tone) is presented. On M3 trials, however, the proportion of time that the comparison process must rely on a biased representation of the test tone would be the critical factor in determining the extent to which the test tone decision is biased by the mask.

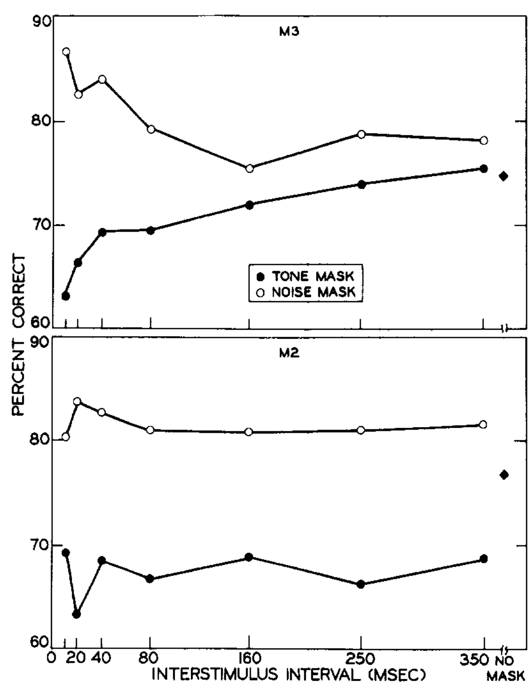

Figure 8. The percentage of correct identifications of the target tones as a function of the silent interstimulus interval for Group S in Experiment 3. (The upper panel illustrates performance on M3 trials; the lower panel illustrates performance on M2 trials.)

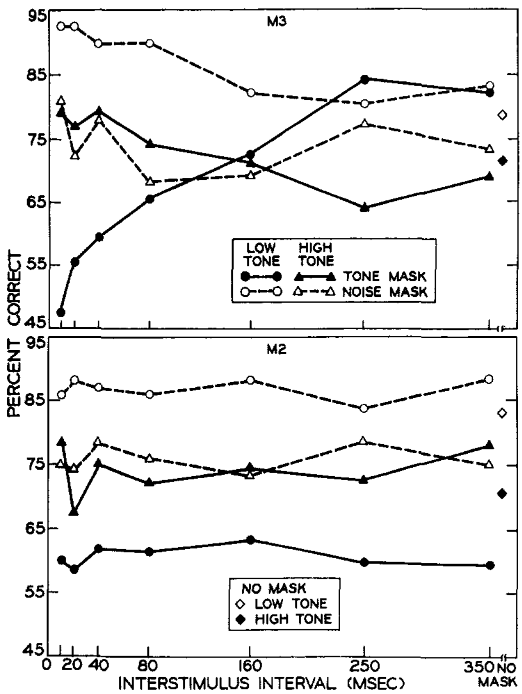

Figure 9. The percentage of correct identifications of the target tones as a function of the silent interstimulus interval for Group S in Experiment 3. (The panels present the data from M2 and M3 trials showing the interaction of target tone and mask type with ISI.)

We thus explain the bias to respond "High" on M2 tone mask trials and on short-ISI, M3 tone mask trials by assuming that the 1,200-Hz tone mask biases the test-tone-synthesized memory toward *high.* But given feedback of *low* on 50% of the trials, subjects may have instituted (on both noise and mask trials), a response criterion bias favoring the low tone to compensate for the *high* bias introduced in synthesized memory on tone mask trials (cf. Hawkins & Presson, 1977). This would explain the result that on both M2 and M3 noise mask trials, there was a bias to respond "Low" that remained constant across ISIs (see Figure 9). We assume that the bias on noise mask trials resulted from the response criterion bias favoring a low response. Similarly, the bias on no-mask trials may be explained by the response criterion bias. And finally, because on M3 trials the effect of the mask on categorization of the test tone would be minimal at the long ISIs, the bias to respond "Low" at the longer ISIs may also be explained by the idea that the response criterion favored a "Low" response.

In sum, the present data on error imbalance are inconsistent with the hypothesis that error imbalance occurs at the perceptual stage of processing, for on M2 trials, error imbalance remained constant across ISIs. This same result supports the idea that error imbalance results from the mask modifying the synthesized memory of the test sound.

The results of an ANOVA conducted on the percentages of correct data for Group N yielded two significant F ratios. There was a significant interaction between day and target, $F(3, 6) = 4.89$, $p < .05$, and the interaction between mask position, day, and target was significant, $F(3, 6) = 4.99$, $p < .05$. Apparently, on M2 trials, subjects developed a bias over days to respond "Low" on M2 trials, whereas on M3 trials a bias to respond "High" developed. However, neither of these two interactions was affected by mask type, $F < 1$ in each case. (For Group S, neither the interaction between target and day nor the interaction between target, day, and mask position was significant ($F < 1$) and $F(3, 15) = 2.13$, $p > .10$, respectively.)

For Group N, neither the main effect of mask type, the interaction between mask type and ISI, nor the three-way interaction between mask type, ISI, and mask position was significant ($F < 1$ in each case). Nor was the four-way interaction between target, mask position, mask type, and ISI significant, $F(6, 12) = 1.48$, $p > .25$.

General Discussion

The purpose of the present investigation was to determine whether mask/test tone similarity effects in backward recognition masking can be reliably demonstrated. If so, it was of interest to determine whether perceptual or memory factors are responsible for these effects. Similarity effects were clearly demonstrated in Experiments 1–3, but only the data from approximately 60%

of the subjects showed these effects. We tentatively suggest below that the individual differences may have reflected different processing strategies employed by different subjects.

Experiment 3 was designed to test whether mask/test tone similarity effects are due to mask interference with a test tone memory or, instead, to selective overwriting of a preperceptual store. The results of Experiment 3 rule out an explanation of similarity effects based on selective overwriting of a preperceptual store.[7] Instead, the data suggest that the mask serves to terminate perceptual processing of the target tone, but, in addition, it interferes with the synthesized auditory memory of the target tone; the extent of this interference depends, in part, on the similarity of the masking tone to the target tone.

Based on the idea that mask interference serves to degrade the test tone comparison process, one might argue that backward recognition masking could conceivably depend entirely on memory interference; in this case, it would not be necessary to postulate a single-channel preperceptual auditory store. However, if backward recognition masking effects depended only on test tone memory interference, backward masking should not have been found on M2 trials in Experiment 3. On these trials, the amount of memory interference prior to presentation of the test tone was held constant across ISIs; consequently, if backward masking depended only on memory interference, the extent of masking should have remained constant across ISIs, which it did not according to the Δf analysis.

A study by Massaro and Idson (1977) also provides evidence against the idea that backward masking effects are based entirely on memory interference. In that study, subjects heard two short tones separated by a variable ISI. The subject's task was to determine whether the second tone was higher or lower in pitch than the first. Note that in this case the second tone served as a comparison tone but also acted to mask the first tone. Performance on the experimental task improved with increases in the ISI that separated the two comparison tones. If backward masking was based entirely on memory interference from the mask, performance should have remained constant across ISIs because the amount of interference with the test tone memory would have remained constant across ISIs.

The results of Experiment 3 reported here along with those of Massaro and Idson (1977) strongly support the view that backward recognition masking effects result from termination of perceptual processing of a target sound's representation held in a preperceptual store. Additionally, test tone memory interference may serve to accentuate backward masking effects, although the individual differences found in the present experiments suggest that interference need not always play a role. Massaro (1970c) found that interference in synthesized auditory memory depends, in part, on the strategies employed by subjects when performing an experimental task. The degree to which different stimuli interfere with information in synthesized memory may depend on the subject's processing strategies. The individual differences found in the present experiments may have resulted from the use of different strategies by different subjects, although at present we have no direct evidence to support this view.

The results of Experiments 1 and 3 suggest that Hawkins' (Hawkins et al., 1974; Hawkins & Presson, 1977) additive model of error imbalance needs some revision.

[7] One possibility suggested by a reviewer is that the flat similarity effect functions found for M2 trials in Experiment 3 may be due to a combination of selective overwriting *and* memory interference. However, in order for this to yield the flat functions found on the M2 trials (see Figure 8), one would have to assume that the effect of similarity due to selective overwriting of the preperceptual image decreased with increases in ISI, but the effect of similarity due to mask interference *increased* with increases in ISI. It would further have to be assumed that the increases in the similarity effects due to interference were exactly complementary to the decreases due to partial overwriting. Although such an explanation is possible, it is not parsimonious, and there is no evidence that the interference effects should increase in this manner with increases in ISI.

Experiment 1 showed that the degree of error imbalance does not increase with increases in the frequency separation between test tone and mask. And Experiment 3 demonstrated that the extent of biasing by the mask does not necessarily depend on the mask's temporal proximity to the target tone. The present data indicate that error imbalance does not result from an integration of target tone and mask percepts. Rather, it appears that the error imbalance effect is most probably based on some sort of biasing of the target-tone-synthesized memory. Further research is needed to clarify this.

The present investigation makes clear the need to determine the stages of processing responsible for effects found in backward recognition masking experiments. We were able to show that target/mask similarity effects and error imbalance do not result from operations at the initial stage of perceptual processing. Target/mask similarity effects and error imbalance have been demonstrated in visual backward masking experiments (Bernstein, Fisicaro, & Fox, 1976; Hellige et al., 1977; Smith, Haviland, Reder, Brownell, & Adams, 1976), and such findings could profit from a similar attempt to isolate the relative contributions of perceptual and memory factors. Analogous to the logic of the present experiments, the results do not *necessarily* indicate that the effects are due to the perceptual stage of visual processing. Visual preperceptual and synthesized memories have been documented (Bongartz & Scheerer, 1976; Massaro, 1975b; Phillips, 1974), and it is not unreasonable to expect that their roles in visual backward masking be similar to those of their counterparts in auditory processing. If this expectation is confirmed, it will augment the evidence for analogous operations in auditory and visual information processing.

Reference Note

1. Kallman, H. J., & Massaro, D. W. *Similarity effects in backward masking* (Rep. No. 4). Madison: University of Wisconsin Human Information Processing Program (WHIPP), 1978.

References

Bernstein, I. H., Fisicaro, S. A., & Fox, J. A. Metacontrast suppression and criterion content: A discriminant function analysis. *Perception & Psychophysics*, 1976, *20*, 198–204.

Bongartz, W., & Scheerer, E. Two visual stores and two processing operations in tachistoscopic partial report. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 203–219.

Coltheart, M., & Arthur, B. Evidence for an integration theory of visual masking. *Quarterly Journal of Experimental Psychology*, 1972, *24*, 262–269.

Deutsch, D. Generality of interference by tonal stimuli in recognition memory for pitch. *Quarterly Journal of Experimental Psychology*, 1974, *26*, 229–234.

Hawkins, H. L., & Presson, J. C. Masking and preperceptual selectivity in auditory recognition. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, N.J.: Erlbaum, 1977.

Hawkins, H. L., Thomas, G. B., Presson, J. C., Cozic, A., & Brookmire, D. Precategorical selective attention and tonal specificity in auditory recognition. *Journal of Experimental Psychology*, 1974, *103*, 530–538.

Hellige, J. B., Walsh, D. A., Lawrence, U. W., & Cox, P. J. The importance of figural relationships between target and mask. *Perception & Psychophysics*, 1977, *21*, 285–286.

Holding, D., Loeb, M., & Yoder, D. "Masking" versus interference in pitch perception. *Journal of Auditory Research*, 1972, *12*, 247–254.

Idson, W. L., & Massaro, D. W. Perceptual processing and experience of auditory duration. *Sensory Processes*, 1977, *1*, 316–337.

Loeb, M., & Holding, D. H. Backward interference by tones or noise in pitch perception as a function of practice. *Perception & Psychophysics*, 1975, *18*, 205–208.

Massaro, D. W. Forgetting: Interference or decay? *Journal of Experimental Psychology*, 1970, *83*, 238–243. (a)

Massaro, D. W. Preperceptual auditory images. *Journal of Experimental Psychology*, 1970, *85*, 411–417. (b)

Massaro, D. W. Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1970, *83*, 32–39. (c)

Massaro, D. W. Preperceptual images, processing time and perceptual units in auditory perception. *Psychological Review*, 1972, *79*, 124–145.

Massaro, D. W. Backward recognition masking. *Journal of the Acoustical Society of America*, 1975, *58*, 1059–1065. (a)

Massaro, D. W. *Experimental psychology and information processing*. Chicago: Rand McNally, 1975. (b)

Massaro, D. W. (Ed.). *Understanding language: An information processing analysis of speech perception, reading, and psycholinguistics*. New York: Academic Press, 1975. (c)

Massaro, D. W. Auditory information processing. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes.* Hillsdale, N.J.: Erlbaum, 1976.

Massaro, D. W., & Cohen, M. M. Preperceptual auditory storage in speech recognition. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception.* Berlin: Springer-Verlag, 1975.

Massaro, D. W., Cohen, M. M., & Idson, W. L. Recognition masking of auditory lateralization and pitch judgments. *Journal of the Acoustical Society of America,* 1976, *59,* 434–441.

Massaro, D. W., & Idson, W. L. Temporal course of perceived auditory duration. *Perception & Psychophysics,* 1976, *20,* 331–352.

Massaro, D. W., & Idson, W. L. Backward recognition masking in relative pitch judgments. *Perceptual and Motor Skills,* 1977, *45,* 87–97.

Moore, J. J., & Massaro, D. W. Attention and processing capacity in auditory recognition. *Journal of Experimental Psychology,* 1973, *99,* 49–54.

Phillips, W. A. On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics,* 1974, *16,* 283–290.

Pollack, I. Identification of random auditory waveforms. III. Effects of interference. *Journal of the Acoustical Society of America,* 1976, *60,* 680–686.

Smith, E. E., Haviland, S. E., Reder, L. M., Brownell, H., & Adams, N. When preparation fails: Disruptive effects of prior information on perceptual recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, *2,* 151–161.

Sparks, D. W. Temporal recognition masking—or interference? *Journal of the Acoustical Society of America,* 1976, *60,* 1347–1453.

Turvey, M. T. On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review,* 1973, *80,* 1–52.

Received April 24, 1978 ■