Perceptual and Motor Skills, 1977, 45, 87-97.

BACKWARD RECOGNITION MASKING IN RELATIVE PITCH JUDGMENTS[1]

DOMINIC W. MASSARO AND WENDY L. IDSON
University of Wisconsin

Summary.—Backward recognition masking refers to interference of a second masking tone with recognition of a target tone presented earlier in time. The degree of interference has been found to decrease as the length of the silent interval separating the two tones increases. These results have been interpreted as representing interference of the masking tone upon the preperceptual storage and perceptual resolution of the target. It is logically possible, however, that the masking tone does not interfere with perceptual resolution but interferes with comparison of the target to a long-term memory representation. The current research was designed to provide a critical test of this alternative hypothesis by modifying the backward recognition masking task. Subjects determined whether the masking tone was higher or lower in pitch than the target tone. The frequencies of the target and masking tones varied randomly across trials. This ensured that the task could not be performed by comparing the target to a representation in long-term memory. Nevertheless, masking was obtained in this task, arguing against the comparison argument and in favor of the perceptual resolution interpretation. Given that masking was obtained under both ipsilateral and contralateral presentation of the tones, the results argue for a central preperceptual auditory storage that holds information after the inputs from the two ears are combined in the auditory system.

An auditory stimulus exists along a temporal dimension. As perception is not immediate, a problem arises as to the nature of perception of a stimulus extended in time. One view of auditory recognition (Massaro, 1975b) argues that the perceptual process itself requires time. According to this theory, the information contained in a stimulus—in this case a pure tone—is held in an early preperceptual storage. During the recognition process, information is extracted continuously from the tone, over a period of approximately 250 msec. If the duration of the tone is less than a quarter of a second, processing will continue after test-tone offset, based upon the sensory representation of the tone.

Some predictions of the model have been supported utilizing a backward recognition masking paradigm. In a typical backward recognition masking experiment (Massaro, 1970), a short target tone is presented, followed after a variable silent intertone interval, by a masking tone. The target is usually one of two fixed alternatives and the subjects' task is to determine which tone was presented on each trial. The finding in these studies is that the masking

[1]This research was supported by Public Health Service Grant MH-19399 and was conducted while the second author was supported by a University of Wisconsin graduate traineeship. Requests for reprints should be addressed to Dominic W. Massaro, Department of Psychology, University of Wisconsin, Charter at Johnson St., Madison, Wisconsin 53706.

tone interferes with identification of the target, the degree of interference decreasing with increases in the silent interval between the two tones. Performance in the backward recognition masking task is usually a negatively accelerated monotonic function of the intertone interval. In terms of Massaro's model (1975b), this masking function represents the extraction over time of a fixed proportion of the information that remains unprocessed. Storage of the masking tone disrupts the preperceptual storage of the target and terminates processing of it.

The intent of the current research is to provide a critical test of one of the premises of this interpretation of backward recognition masking. The assumption has been made that the masking tone interferes with the sensory representation of the target in preperceptual store, terminating processing at this level of analysis. However, another interpretation of these results is tenable. Performance in the backward masking task reflects two sequential processes. The pitch of the target tone must initially be resolved from the information in preperceptual store. In order to make an absolute identification of the target, this percept has to be compared to the long-term memory representations of the two alternative tones. Logically the interference exercised by the mask could occur either during pitch resolution or during the following comparison process. Either interpretation would be consistent with the improvement in performance which accompanies increases in the silent intertone interval.

This distinction is crucial as the two alternatives localize the effects of backward recognition masking at different stages of processing. Under the assumption that the mask interferes with perceptual resolution of the target, the results of backward masking studies have been used to explore the temporal course of perceptual processing of a variety of perceptual attributes, such as pitch (Hawkins, Thomas, Presson, Cozic, & Brookmire, 1974; Massaro, 1970, 1975a; Watson, Wroton, Kelly, & Benbassat, 1975), timbre (Massaro, 1972), loudness (Moore & Massaro, 1973), lateralization (Massaro, Cohen, & Idson, 1976), duration (Massaro & Idson, 1976), and consonant and vowel quality (Massaro, 1975c; Wolf, 1976). However, if the mask functions as an interference tone at the memory-comparison stage, a perceptual interpretation of the backward recognition masking results is unjustified.

Exp. 1 was designed to distinguish between these alternatives. In the usual backward masking task, the frequencies of the target tones were fixed throughout the experiment. As a result subjects had to maintain and utilize long-term memory representations of the target tones in performing the task. Consequently these previous studies cannot differentiate between the alternative explanations of backward recognition masking. In the current study, the backward recognition masking task was modified to preclude any possible comparison with a long-term memory representation of the target. If backward masking can be demonstrated in this task, the conclusion can be drawn that long-term memory comparison is not a necessary condition for eliciting backward recognition masking. This would in turn support the interpretation that backward recognition masking results from interference with the target information in preperceptual storage.

In most backward recognition masking tasks the subject is asked to make an absolute identification of the target. In Exp. 1, the task was altered to require a relative judgment between the two tones. The subjects' task was to determine on each trial whether the masking tone was higher or lower in pitch than the target tone. Roving frequencies were employed for both the target and masking tones. On each trial the frequency of the target tone was chosen randomly, within a range of 240 Hz, and the frequency of the masking tone computed with reference to it. This would prevent the subjects from learning to label the sounds and making their comparison on the basis of the labels (cf. Massaro, 1975b, pp. 479-481). To ensure that subjects were not simply listening for a specific frequency difference between the tones, two directions (higher or lower), and four extents of frequency difference were employed: 5, 10, 20, and 30 Hz. The wide range of values from which the target frequency was chosen made it impossible for the task to be performed by comparing the target to a long-term memory representation. Rather the pitch of the target would have to be resolved and compared to the pitch of the mask in order to determine the relationship between the two tones. Unlimited time would exist for processing the masking tone, since it is the last tone in the two-tone sequence. Accordingly the masking tone should be resolved equally well under all intertone intervals and any change in performance with intertone interval cannot be due to differences in perception of the masking tone. Furthermore the masking tone cannot be said to interfere with the comparison process since it is one of the tones to be compared. It follows that the intertone interval should not affect the resolution of the masking tone or the operation of the comparison process but must have its effect via interference of the masking tone on resolution of the pitch of the target tone.

If masking comparable to that obtained in an absolute recognition masking task with fixed target frequencies is found in Exp. 1, it will argue strongly that the mask is interfering with perception of the target. Performance would be a direct function of the amount of processing time available between the onset of the target and masking tones. The most parsimonious interpretation of this result and the findings in other backward recognition masking studies would be that presentation of the mask terminates further resolution of the pitch of the target. As a by-product, demonstrating masking in this context, which requires a relative judgment between tones, would provide another demonstration that the phenomenon generalizes beyond an absolute identification task.

EXPERIMENT 1

Method

Subjects.—The subjects were eight psychology undergraduates who received credit for their participation.

Stimuli.—The frequency of the target tone was chosen randomly on each trial from the range of 580-820 Hz. The frequency of the masking tone was computed as the frequency of the target tone $\pm$ a variable frequency (Δf) where Δf could assume values of 5, 10, 20, or 30 Hz. The sine wave tones had durations of 20 msec. and were turned on and off at a zero crossing, with an instantaneous rise time. The peak to peak intensity was held constant across frequency changes, at a level equal to 80 dB SPL for a steady-state 650-Hz tone.

All experimental events were controlled by a PDP-8L computer. The tonal stimuli were generated by a digitally controlled oscillator (Wavetek Model 155) and were presented binaurally over matched headphones (Grason-Stadler Model TDH-49). Four subjects were tested simultaneously in separate sound attenuated rooms.

Procedure.—The experiment was conducted in five consecutive days. Each day was divided into two sessions of 370 trials each. All of Day 1 and the first 20 trials of the eight subsequent experimental sessions were treated as practice, though the subjects were unaware of this. On every trial two tones were presented, separated by a variable silent interval of 5, 25, 45, 85, 165, 255, 355, or 505 msec. The subjects' task was to determine whether the second masking tone was higher or lower than the target tone. The subjects had 1.5 sec. from target offset in which to make their response by pressing one of two labeled buttons. Following the response interval .5 sec. of feedback was provided over a visual display of light emitting diodes (Mansanto Model MDA-III). The intertrial interval was 1 sec. All 64 experimental conditions (8 levels of frequency separation $\times$ 8 interstimulus intervals) were completely random and were programmed to occur with equal probability within a session.

Results and Discussion

The dependent measure is the percentage of correct identifications of the frequency relationship between the target and mask. The response frequencies were pooled in the analysis over the four experimental days (Days 2 to 5). Preliminary analysis revealed that the task was too difficult when the two tones were separated by as little as 5 or 10 Hz. Performance was essentially at chance at all interstimulus intervals under these conditions. As a result, the data were reanalyzed with frequency separations of 20 and 30 Hz only.

Fig. 1 presents the percentage of correct identifications of the frequency relationship of the test and masking tones as a function of the silent intertone interval. Performance improved with increases in the intertone interval, within the range of 85 and 255 msec., and remained at this asymptotic level out to the

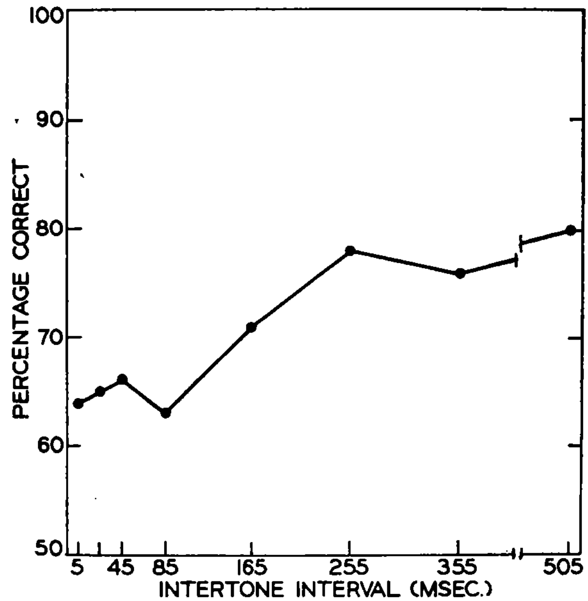

FIG. 1. Percentage of correct identifications of the relationship between the target and masking tone frequencies as a function of the duration of the silent intertone interval (Exp. 1)

longest separation of 505 msec. ($F_{7,49} = 18.17$, $p < .001$). Though differences in over-all performance were found, essentially the same trends were present in the data from individual subjects. Data for four of the subjects are shown in Fig. 2, for a 30-Hz separation between tones.

The consonance of this result with previous work on backward recognition masking argues that the same structures and processes are critical in both tasks. As the subjects could not use a long-term memory representation in performing the present task, the interpretation that the masking tone simply disrupts the comparison of the target tone to a long-term memory representation cannot describe the present results. The interpretation that the mask interferes directly with the preperceptual storage of the target would appear the most reasonable. Insufficient information would be extracted from the target at short interstimulus intervals to allow an accurate identification of the test-mask relationship, despite full processing of the mask. The results of Exp. 1 thus support the premise of Massaro's (1975b) model that masking occurs at the level of preperceptual auditory storage. In addition they demonstrate that masking of a first tone by a second is possible even when a relative judgment is required.

Performance was better at the three shortest intervals than would be expected from a strict monotonically increasing function. At intertone intervals of 5, 25, and 45 msec., subjects identified the relationship between tones slightly

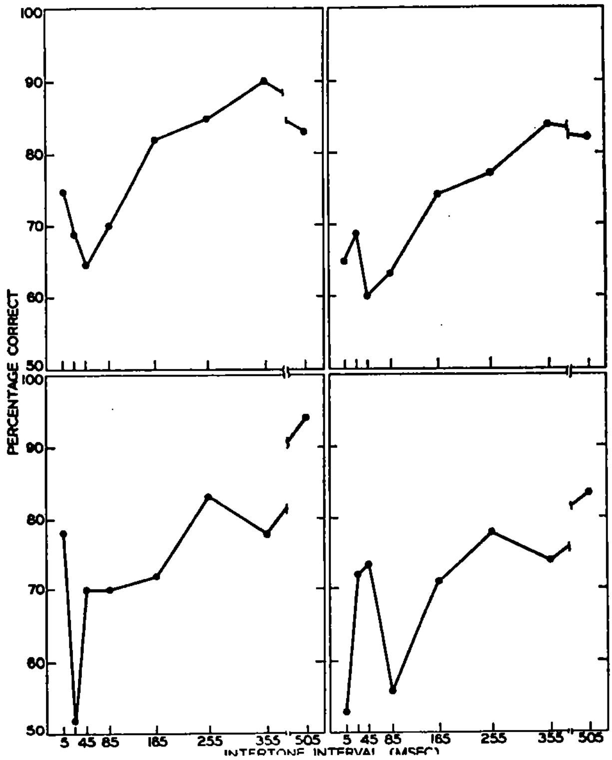

FIG. 2. Percentage of correct identifications of the relationship between the target and masking tone frequencies, with a 30-Hz separation, as a function of the silent intertone interval (four observers in Exp. 1)

more accurately than they did at 85 msec. The slight upturn in performance at the shortest silent intervals has been obtained previously in backward masking studies (Massaro, 1970; Massaro, *et al.*, 1976) and has been taken to represent a period of integration of the information in the target and the mask at a peripheral level (Massaro, 1975a, 1975b).

Exp. 2 was designed to further substantiate the assumption that masking an absolute judgment and a relative judgment involve the same process. In backward recognition masking of an absolute pitch judgment, contralateral presentation of the mask produces results comparable to those found for binaural presentation of the two tones (Massaro, 1970; Massaro & Cohen, 1975). This finding has been taken to indicate that preperceptual auditory storage is located at a central rather than a peripheral level. If relative judgments also utilize information in a central preperceptual store, then contralateral presentation of the tones should also produce masking. Consequently, performance under contralateral and ipsilateral presentation of target and mask was compared in Exp.

2. The subjects were still required to make a relational judgment between the target and masking tones. However, on the ipsilateral trials both tones were presented monaurally to the same ear. On contralateral trials the tones were presented to opposite ears. The ipsilateral condition was included to provide a fair comparison to the contralateral case. Contrasting contralateral, and thus monaural, presentation with binaural presentation would bias the results against obtaining comparable masking in the two conditions, as frequency resolution of a sound is better under binaural than monaural presentation. Consequently, a monaural comparison of contralateral and ipsilateral presentation was employed.

EXPERIMENT 2

Method

Subjects.—The subjects were eight undergraduates, six of whom received credit towards an introductory psychology course for their participation and two of whom received $1.50/hr. for serving in the experiment.

Stimuli.—The frequency of the target tone was chosen randomly on each trial from the range of 550 to 650 Hz. The frequency of the masking tone was computed as the frequency of the target tone $\pm$ Δf, where Δf was 15, 20, 25, or 30 Hz. The two tones on any trial were presented either ipsilaterally or contralaterally. The ear of presentation of the test tone was chosen randomly with equal probability. The output of the oscillator was gated by two computer controlled audio switches (Iconic Model No. 1037) to one of two amplifiers (McIntosh Model MC-50), one for each ear. All other details were the same as in the first experiment.

Procedure.—Two sessions of 320 trials each were conducted on each day for 5 consecutive days, with the exception of Day 2. On Day 2 the subjects received 220 trials/session. All of Day 1 and the first 20 trials of each of the eight experimental sessions were treated as practice though the subjects were unaware of this.

On each trial the subjects heard two test tones separated by a variable silent intertone interval. The subjects' task was to determine whether the masking tone was higher or lower than the test tone. The subjects were given 1.5 sec. from the offset of the second tone in which to make a response. All 256 experimental conditions (8 intertone intervals $\times$ 8 levels of frequency separation $\times$ ipsilateral/contralateral $\times$ ear of target location) were completely random and were programmed to occur with equal probability within a session. All other procedural details were equivalent to Exp. 1.

Results and Discussion

The dependent variable is the percentage of correct identifications of the frequency relationship between the target and masking tones. The response frequencies were pooled over the four experimental days (Days 2 to 5). Fig. 3 presents the percentage of correct identifications of the target-mask relationship, under both contralateral and ipsilateral presentation of the tones, as a function of the duration of the silent intertone interval. It can be seen that performance improved consistently with increases in the duration of the intertone interval. Moreover no appreciable differences resulted from contralateral versus ipsilateral presentation of the tones. The masking functions for the two

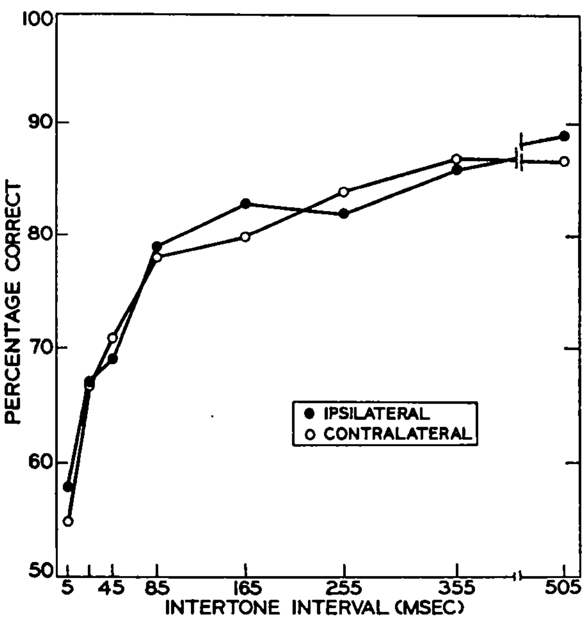

FIG. 3. Percentage of correct identifications of the relationship between the target and masking tone frequencies as a function of the silent intertone interval, under contralateral and ipsilateral presentation (Exp. 2)

modes of presentation are virtually identical. Essentially the same results can be seen in the data for the individual subjects, shown in Fig. 4. Analysis of variance supported these observations. While a significant main effect was found for the intertone interval ($F_{7,49} = 39.85$, $p < .001$), neither the main effect for ipsilateral/contralateral presentation nor the interaction of this variable with intertone interval was significant. Performance improved about 7% with an increase in frequency separation between the two tones ($F_{3,21} = 22.11$) and the masking functions had the same shape for all four Δf values.

The equivalence of performance in the contralateral and ipsilateral presentations supports the idea that preperceptual auditory storage is located centrally after inputs from the two ears are combined in the auditory system. The second

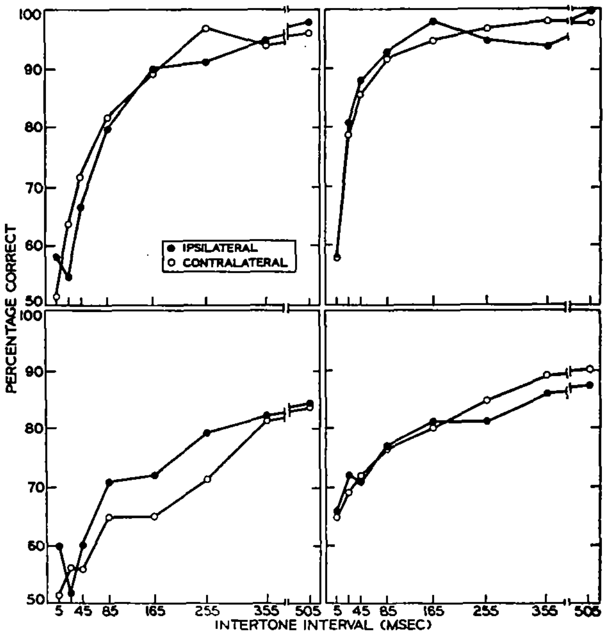

FIG. 4. Percentage of correct identifications of the relationship between the target and masking tone frequencies as a function of the intertone interval, under contralateral and ipsilateral presentation (four observers in Exp. 2)

masking sound interferes with the storage of the earlier test sound even though the sounds are presented to opposite ears. This result also argues that the same operations are producing the masking functions in the absolute judgment and relative judgment tasks. One might argue that the contralateral presentation increases the difficulty of the successive comparison task for some other reason than interference with a preperceptual storage. If this tack is taken, however, the mechanism would have to produce exactly the same results as backward masking, since the ipsilateral and contralateral functions are equivalent.

GENERAL DISCUSSION

The close comparability of the results of Exps. 1 and 2 to those obtained in absolute judgment tasks, employing fixed target frequencies, suggests that recognition masking in a relative judgment involves the same structures and processes as in the absolute judgment. In both cases the target tone is stored in a central preperceptual memory. Resolution of the pitch of this tone occurs during the primary recognition process. If a masking tone is presented before the target tone has been completely processed, the mask will terminate further resolution of the pitch of the target. In the relative judgment task the percept of the target will be compared to the pitch of the mask, while in an absolute judgment task, the percept will be compared to a long-term memory representation of the target alternatives. In both cases, however, the mask appears to interfere with resolution of the target tone.

The present results illuminate previous findings of poorer performance at short interstimulus intervals in successive comparison tasks using short-duration stimuli. Pisoni (1973) found that same-different judgments of 50-msec. vowels improved as the interval between the sounds was increased from 0 to 250 msec. Small and Campbell (1962) and Carbotte (1973) studied successive judgments of the relative durations of short-duration stimuli or empty intervals surrounded by tone bursts. Performance was disrupted if the stimuli were separated by 250 msec. or less relative to that at longer interstimulus intervals. Similarly, Taylor and Smith (1975) observed that shortening the interval to less than 500 msec. in a two-interval forced-choice task decreased amplitude discrimination and detection performance. All of these results support the idea that a second auditory stimulus can interfere with the perceptual processing of a sound presented earlier in time if the second sound occurs within roughly one-fourth of a second after the first sound.

REFERENCES

CARBOTTE, R. M. Retention of time information in forced-choice duration discrimination. *Perception and Psychophysics*, 1973, 14, 440-444.

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.

MASSARO, D. W. Preperceptual auditory images. *Journal of Experimental Psychology*, 1970, 85, 411-417.

MASSARO, D. W. Stimulus information vs. processing time in auditory pattern recognition. *Perception and Psychophysics*, 1972, 12, 50-56.

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. Preperceptual images, processing time, and perceptual units in speech recognition. In D. W. Massaro (Ed.), *Understanding language: an information processing analysis of speech perception, reading, and psycholinguistics*. New York: Academic Press, 1975. Pp. 125-150. (c)

MASSARO, D. W., & COHEN, M. M. Preperceptual auditory storage in speech recognition. In A. S. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception*. New York: Springer-Verlag, 1975. Pp. 226-245.

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.

MOORE, J. J., & MASSARO, D. W. Attention and processing capacity in auditory recognition. *Journal of Experimental Psychology*, 1973, 99, 49-54.

PISONI, D. B. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 1973, 13, 253-260.

SMALL, A. M., & CAMPBELL, R. A. Temporal differential sensitivity for auditory stimuli. *American Journal of Psychology*, 1962, 75, 401-410.

TAYLOR, M. M., & SMITH, S. M. Monaural detection with contralateral cue: V. Interstimulus interval in MDCC and amplitude discrimination. *Journal of the Acoustical Society of America*, 1975, 57, 1500-1511.

WATSON, C. S., WROTON, H. W., KELLY, W. J., & BENBASSAT, C. A. Factors in the discrimination of tonal patterns: I. Component frequency, temporal position, and silent intervals. *Journal of the Acoustical Society of America*, 1975, 57, 1175-1185.

WOLF, C. G. A recognition masking study of consonant processing. *Perception & Psychophysics*, 1976, 19, 35-46.

Accepted May 12, 1977.