Backward interference by tones or noise in pitch perception as a function of practice

MICHEL LOEB and DENNIS H. HOLDING
University of Louisville, Louisville, Kentucky 40208

In a first experiment, the identification of brief tonal stimuli was unaffected by subsequent presentation of white noise. However, subsequent pure tones, whether central to the stimulus frequencies or remote from them, caused substantial declines in correct identification. Apparent backward masking seems therefore to follow rather than to precede some degree of categorization of the masking stimulus. A second experiment shows that even these effects are temporary. Early masking effects are heavily modified by practice, and are not therefore permanent features of sensory processing. Neither experiment provides support for preperceptual theory.

It has been suggested by Massaro (1972) that there exists a basic difference between the suffix effect, in "precategorical acoustic storage," and the backward masking evidence for "preperceptual auditory images." Although Crowder and Morton (1969) regard the interference generated by a redundant suffix as satisfactory evidence of precategorical storage, the amount of interference generated depends upon the similarity of the suffix to the stimulus materials. In particular, neither white noise nor a buzzer give rise to interference.

In contrast, backward recognition masking is viewed as not affected by similarity. Massaro (1970) had shown that varying the similarity between masking and test tones over a range of approximately two to one did not affect the amount of interference, while indirect evidence was available (Dirks & Bower, 1970) to show in addition that white noise interferes retroactively with speech intelligibility. However, subsequent work (Holding, Loeb, & Yoder, 1972) has provided a direct comparison between tonal and noise masks, showing that in some experimental conditions the interference generated by tones greatly exceeds the amount due to noise masks. It therefore seems dubious whether the recognition masking paradigm provides any better evidence for preperceptual processes than does the suffix effect.

The comparison made by Holding et al. (1972) was based upon a two-tone task (Elliott, 1970) in which the subject has to compare and discriminate a second tone from the first. Since the Massaro (1970) data were obtained from a one-tone task, in which the

subject is required to identify the pitch of a single tone, it seems desirable to obtain a direct comparison of the masking effects of tone and noise in a similar task. The two paradigms should yield similar results, since it has been shown (Loeb & Holding, 1972) that subjects tend to deal with the two-tone task by making absolute identifications of whichever single tone is the variable one.

The first experiment was designed to verify the predicted correspondence, comparing tone and noise masks in a one-tone identification procedure. In addition, two tonal masks of widely differing pitch were compared, as a check on the generality of the earlier data.

EXPERIMENT I

Method

Five volunteer subjects were used. The basic task was to identify as "high" or "low" pure tones of 770 or 870 Hz, which were presented randomly in blocks of 50 tones. Presentation of a stimulus tone was always preceded by a 1-sec warning light initiated 2 sec before the tone.

In the control condition, the tones were presented binaurally via earphones for 20 msec at 81 dBA. Rise and decay times were 2.5 msec. In the three experimental conditions, these were followed by a 500-msec masking sound at 81 dBA, after a 40-msec silent interval. In the "noise" condition, the masking sound was band-limited (100-10 kHz) thermal noise produced by feeding the output of a Grason-Stadler noise generator through a Krohn-Hite Model 350D filter; the "central tone" masking condition used a pure tone at 820 Hz; the "remote" condition employed a 1,689-Hz pure tone. Subjectively, the loudness of tones and noise at 81 dBA is approximately equivalent, although the broad bandwidths employed for the noise should make for a slightly higher loudness level (Zwicker, Flottorp, & Stevens, 1957).

All subjects attended on 3 separate days for practice sessions, and on 4 subsequent days for the test sessions which yielded the scored data. Each practice session comprised 100 control trials followed by 200 trials in one of the experimental conditions. Each test session was preceded by 10 warm-up trials and consisted of 200 trials, 50 on each condition, run in a counterbalanced order across the 4 test days.
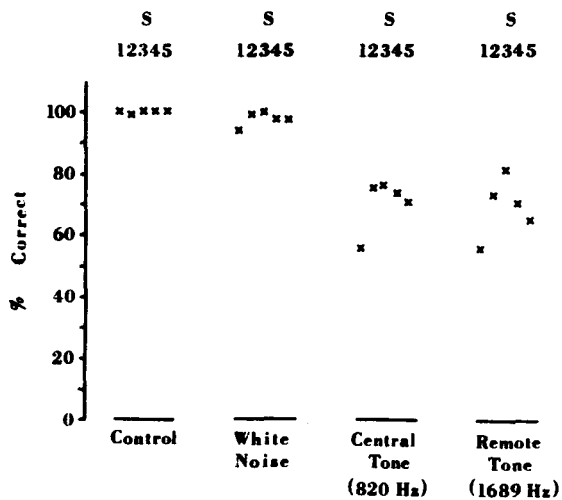
Figure 1. Percentage correct identification of 770- and 870-Hz tones by each subject in control and masking conditions.

Results

The data indicate clearly that white noise produces no backward masking, but that both central and remote pure tones give rise to equal and substantial effects. There were no differences between identification of the high and low stimulus tones. The percentages of total correct identifications are shown, separately for each subject, in Figure 1.

The obviousness of the results probably renders analysis superfluous. However, as a matter of form, the arcsin transforms of the subject percentages were subjected to analysis of variance. The difference between conditions is highly significant (F = 23.4, df = 3,1. p < .001). No further analysis appeared necessary.

EXPERIMENT II

The results of Experiment I pose some difficulty for preperceptual or precategorical theories. In experimental conditions comparable to those of Massaro (1970), white noise shows no backward masking effect, despite the fact that the range of frequencies in the noise bandwidth includes the frequencies of the stimulus tones. In contrast, the remote pure tone has as great an effect as the central tone, although far removed from the stimulus frequency, presumably because it is perceived as qualitatively similar. The difference can only reflect the perceptual category to which the subject assigns the tone or noise before the interference effect takes place. If, instead, it is postulated that a preperceptual screening process is applied to the physical characteristics of the mask, the theory has to account for a fairly complex form of preperceptual categorization.

However, it further appears that backward masking effects are not a permanent feature of auditory

perception. Leshowitz and Cudahy (1973) found their listeners discriminating differences of 15 Hz in conditions resembling those where Massaro's (1970) subjects were responding at chance levels to 100-Hz differences, ascribing the discrepancy to practice effects. The data of Experiment I also suggest that considerable practice effects exist. Since no learning curves appear to be available, Experiment II was designed to exhibit the effects of practice under tonal and noise masking conditions. Successive difference limens were obtained in this experiment, in preference to the percentages of correct identifications used by Massaro (1970) and hence in Experiment I.

Method

Six new subjects volunteered for extended testing. The basic task remained the same as in Experiment I. The masking conditions also remained the same, except that the central tone mask was eliminated because the threshold procedure made its location ambiguous. The three remaining conditions thus consisted of remote tone mask, noise mask, and control.

The first two subjects were tested for 6 and 11 days, respectively. One stimulus was fixed at 820 Hz, while the other began at 156 Hz higher (tone mask) or 19 Hz higher (control and noise mask). The subjects were tested at these separations for 50 trials, or for 10 consecutive correct responses. Testing continued according to an adaptive procedure resembling the PEST technique (Taylor & Creelman, 1967), with the upper frequency descending in constant ratio (0.8) steps, in 50-trial blocks (or 10 correct) until performance became less than 75% correct. Alternating ascending and descending steps were next used to bracket the difference limen. The procedure entailed approximately 350-400 trials per session, only one session taking place on any experimental day. The data for these subjects is virtually the same as for the later subjects, and both sets of data are therefore presented.

The remaining four subjects were each tested once per day for 20 days. For these subjects, the difference limens were obtained more efficiently by using 20-trial blocks at each step size (unless the first 10 trials were all correct). Progression continued downward, in 0.8 steps, from an estimated starting point for each session, until accuracy fell below 75%. If accuracy at that stage was greater than 60%, the D.L. was interpolated; if accuracy fell below 60%, the upper tone was reset at one step size below the starting point for that session. Each subject required a mean of 190 trials per session, or 3,800 trials in all. For all six subjects, limens were obtained for all three conditions on each experimental day, in a counterbalanced order.

Results

All six subjects improved in all three conditions during the testing period. Of course, six subjects showing the same trend is sufficient for significance on a two-tailed binomial test, p < .05. The practice curves for individual subjects are given in Figure 2, in terms of means for each block of four sessions. Comparing these curves reveals some individual differences. Although subjects K.K. and M.K. make improvements with practice, neither set of masking curves converges with the control condition, the discrepancy remaining anomalously large for subject K.K. However, the remaining subjects (M.L., W.L., F.F., and D.O.) show virtual convergence. To illustrate the changes from session to session, the
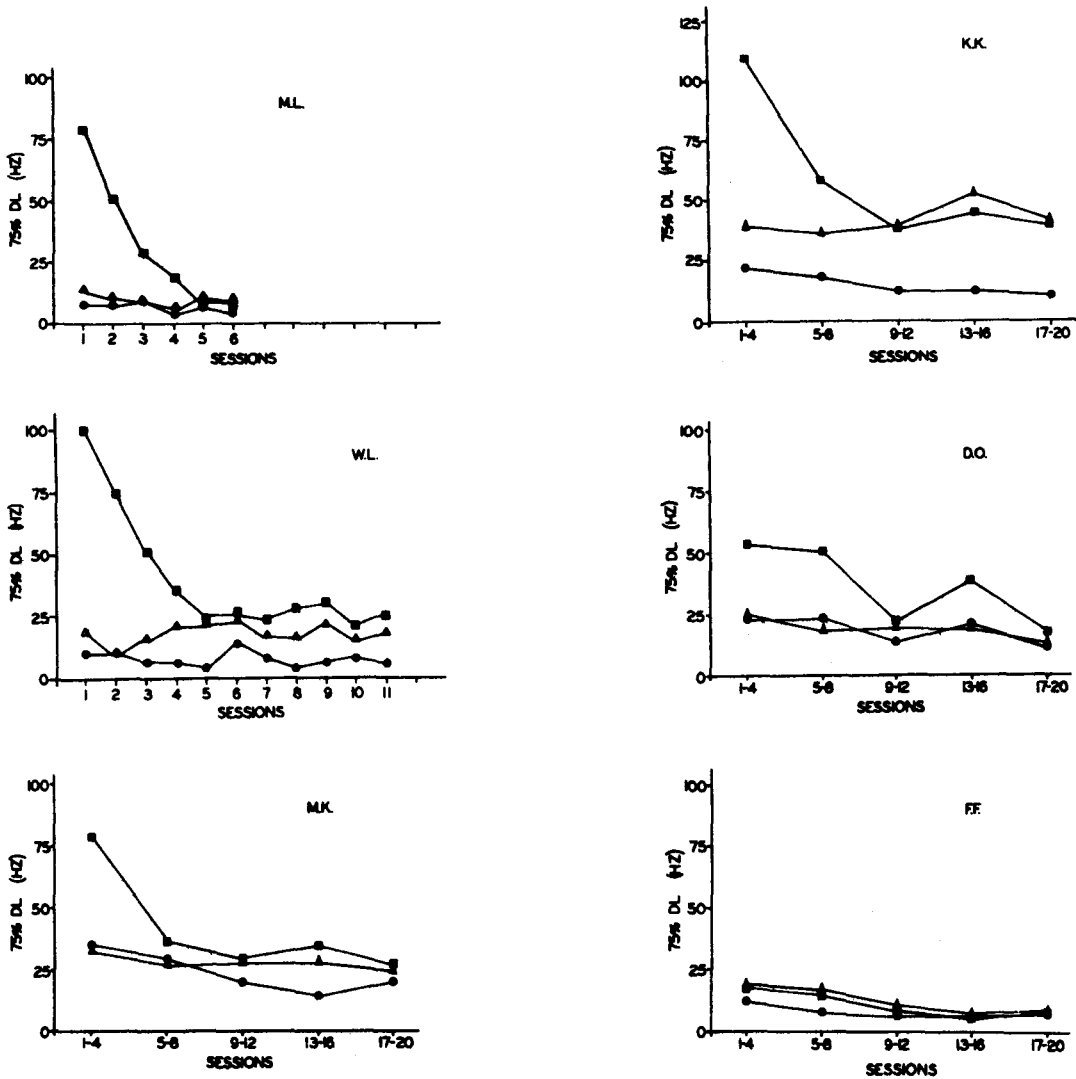
Figure 2. Practice curves for six individual subjects.

means averaged across subjects are plotted in Figure 3. Clearly, the tonal mask condition yields the largest initial limens, confirming the results of Experiment I, but this condition also shows the greatest practice effect. In general, both tone and noise mask conditions tend to asymptote at levels only a few hertz removed from the control condition.

DISCUSSION

The relationship between the amounts of masking by tones or by noise in Experiment I are not indicative of direct interference at a sensory level. Since concurrent white noise would normally produce more impairment than a remote tone, the lack of resemblance to direct acoustical masking is such as to suggest that the term "masking" is a misnomer in this
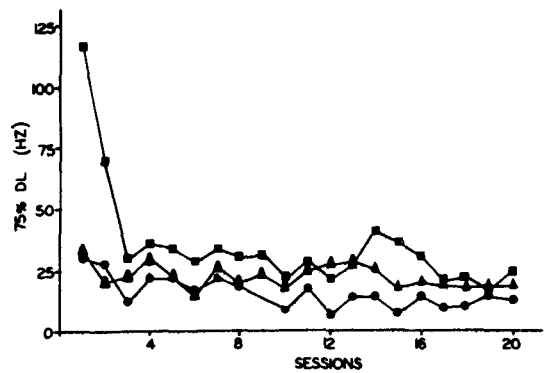

Figure 3. Mean limens as a function of practice in three masking conditions.

context. A similar effect has been noted at the level of speech identification, where Darwin (1971) has found that the recognition of stop-vowel syllables is more impaired by dichotic presentation of a different stop-vowel syllable than by the same vowel, a different vowel, or a nonspeech timbre. Thus, in the cases of both simple pitch identification and of complex speech recognition, the interference effect appears to be to some extent postcategorical. It is difficult to determine to what extent this is true, since the preperceptual and precategorical hypotheses are insufficiently explicit. Quite possibly, the notion of a sharp dichotomy between perceptual and pre-perceptual processing will be better replaced by an analysis assuming various levels of processing (cf. Craik & Lockhart, 1972).

The results of Experiment II bear out the suggestion made by Leshowitz and Cudahy (1973). The backward masking or interference phenomenon, upon which the Massaro (1970) theory is based, is apparently susceptible to a major practice effect. Most subjects can learn to disregard the interference initially caused by the backward masker. A learning effect of this kind suggests that the interference, when observed in naive subjects, must take place at a level which is relatively more central than the peripheral level implied by the preperceptual hypothesis. Despite the small, residual differences between masking conditions, it seems clear that backward interference is not a stable feature of practiced performance. The effect cannot, therefore, play any major part in the construction of a preperceptual theory.

REFERENCES

CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Behavior*, 1972, **11**, 671-684.

CROWDER, G. G., & MORTON, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, **5**, 365-373.

DARWIN, C. J. Dichotic backward masking of complex sounds. *Quarterly Journal of Experimental Psychology*, 1971, **23**, 386-392.

DIRKS, D. D., & BOWER, D. Effect of forward and backward masking on speech intelligibility. *Journal of the Acoustical Society of America*, 1970, **47**, 1003-1008.

ELLIOTT, L. L. Pitch memory for short tones. *Perception & Psychophysics*, 1970, **8**, 379-384.

HOLDING, D. H., LOEB, M., & YODER, D. "Masking" versus interference in pitch perception. *Journal of Auditory Research*, 1972, **12**, 247-254.

LESHOWITZ, B., & CUDAHY, E. Frequency discrimination in the presence of another tone. *Journal of the Acoustical Society of America*, 1973, **54**, 882-887.

LOEB, M., & HOLDING, D. H. Delayed interference in pitch judgments. *Journal of Auditory Research*, 1972, **12**, 336-339.

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

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

TAYLOR, M. M., & CREELMAN, C. D. PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 1967, **41**, 782-787.

ZWICKER, E., FLOTTORP, G., & STEVENS, S. S. Critical bandwidth in loudness summation. *Journal of the Acoustical Society of America*, 1957, **29**, 548-557.