Auditory backward recognition masking: Effect of interaural phase on masker efficacy

R. D. SHILLING and D. R. SODERQUIST

University of North Carolina, Greensboro, North Carolina

The effect of interaural phase on pitch and lateralization recognition was examined. Tonal signals were followed by a variable interstimulus interval and an interference tone. Stimuli were presented in a binaural masking-level difference paradigm as a means of manipulating the perceptual location of the signal within the head. In separate tasks, subjects were required to recognize the pitch or location of the signal. The pitch-recognition task resulted in the expected increase in performance as the interstimulus interval increased. There was no effect of interaural phase on pitch-recognition performance. There was no significant effect of interstimulus interval on performance in location recognition. Subjects were proficient at recognizing location of the signal regardless of the interstimulus interval. The data suggest that a strict single-channel interpretation of the preperceptual storage model of backward auditory recognition is inadequate.

The functional processes that underlie auditory perception have been investigated with a variety of experimental paradigms. Auditory recognition, for example, has been assumed to emerge from the sequential processing of stimulus attributes. The stimulus waveform contains the information required to yield simple percepts such as pitch, timbre, loudness, location, and duration. In this context, Massaro (1970, 1972) and his colleagues (Kallman & Massaro, 1979, 1983) have suggested a model that specifies a centrally located *preperceptual auditory store* (PAS) and a *synthesized auditory memory* (SAM). The PAS is hypothesized to store acoustic features while simultaneously transmitting them to SAM. The features stored in the PAS decay in 250 to 300 msec and are subject to being overwritten. Thus, information needed to synthesize an adequate percept may be unavailable due to decay or overwriting. It is assumed that such disruptions lead to the degradation of recognition and to an incomplete synthesis.

One way to examine the process of information disruption is to use an *auditory backward recognition-masking* (ABRM) stimulus configuration. In such a paradigm, the signal to be recognized is followed in each interval by a silent period (an interstimulus interval, ISI) and a subsequent interference tone. By varying the ISI, one may

determine the temporal interval over which degradation occurs. This procedure, and others, has yielded the aforementioned 250–300 msec interference interval. Short ISIs yield poor (near chance) performance, because, according to the model, the information stored in the PAS is overwritten before the SAM receives a complete representation of the stimulus.

The investigations that have attempted to delineate the scope of auditory recognition have been as controversial as they have been diverse. Differences in data interpretation, stimulus parameters, methodological controls, and experimental paradigms have all contributed to the debate. For example, Yost, Berg, and Thomas (1976) showed that the psychophysical procedure one uses affects the experimental outcome. Furthermore, Loeb and Holding (1975) and Leshowitz and Cudahy (1973) suggested that the training and experience of the subject were critical parameters. Watson and his colleagues (Watson, Wroton, Kelley, & Benbassat, 1975) revealed once again that as stimulus uncertainty increases performance decreases, thus augmenting the ABRM effect. Massaro (1975), in contrast, has suggested a different perspective and reported supporting evidence across different stimulus configurations and psychophysical techniques.

One particular psychophysical paradigm, two-interval forced choice (2IFC), lends itself to the examination of another important theoretical issue in auditory recognition. This question concerns the perceptual independence of acoustic features assumed to be stored in the hypothesized PAS and synthesized by a separate SAM. If stimulus features are centrally stored within a single "cluster" (i.e., in a single channel), then the recognition of two different percepts, location and pitch, should be equally degraded by a subsequent interference tone. In sum, any disruption, at short ISIs, should affect the stimulus features underlying both location and pitch. On the other hand, subjects may use a variety of cues to improve per-

This research was supported by grants to D. R. Soderquist from the National Institute of Neurological and Communicative Disorders and Stroke (1 R15 NS23187-01A1) and a Research Council Grant from the Graduate School of the University of North Carolina at Greensboro. Portions of this paper were based on a study by R. D. Shilling as partial fulfillment of requirements for the MA degree at the University of North Carolina at Greensboro. The data were presented at the 1989 Mid-winter Meeting of the Association for Research in Otolaryngology. The authors wish to express their special thanks to Aloha E. Bryson for her assistance with statistical analysis. Reprint requests may be addressed to R. D. Shilling, Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27412, or via electronic mail: SHILLING@UNCG.BITNET.

formance. Thus, it is just as logical to conclude that auditory recognition would show an improvement if the subject were to use another percept (e.g., location) as a means of separating the signal (i.e., pitch) from the interference tone. The latter possibility would occur if the features necessary for location and pitch were available at the same time and stored separately. In the present experiments, we examined these questions by combining the 2IFC and ABRM paradigms with a *binaural masking-level difference* (BMLD) stimulus configuration.

The independent manipulation of the location percept can be accomplished with the well-known BMLD stimulus configuration. In the presence of a relatively intense diotic background noise, low-frequency tonal stimuli are lateralized (localized) differentially within a subject's head by interaural phase manipulations. For example, when the signals (e.g., 430 and 450 Hz) have an interaural phase of zero, and the interference tone (e.g., 440 Hz) has an interaural phase of 180° , the signals will be perceived at midline, and the interference tone will be perceived as lateralized toward the ears. The performance of the subjects in these homophasic and antiphase conditions can be used to examine questions about the recognition of pitch and localization.

METHOD

Subjects

Five adult normal-hearing subjects participated in the experiments. Two of the subjects completed all experimental conditions (F.D. and S.M.) and were paid for their services. The other subjects were the authors and a graduate student volunteer.

Apparatus and Stimuli

Each subject was seated in a double-walled sound-attenuated chamber and listened to binaural stimuli through matched earphones (TDH-49). The latter were mounted in supraural cushions (MX41/AR). A response panel with lights and buttons was used to indicate when to listen and respond as well as to provide feedback. The signals and interference tone were digitally generated, using ILS-PC software on an IBM AT microcomputer. The 20-msec stimuli had a 5-msec linear rise/fall and a 10-msec steady state. The stimuli were converted by a Data Translations DT2801A D/A converter at a 25-kHz sampling frequency. Analog noise, produced by a Hewlett-Packard 8057-A precision noise generator, was bandpassed (10 Hz to 3 kHz, 3 dB down) through a 48 dB/Oct Coulbourn Instruments S75-36 adjustable filter. The stimuli were mixed, amplified, and controlled with a Coulbourn Instruments mixer/amplifier (S82-24) and programmable attenuators (S85-08). The stimuli were monitored with a Tektronix storage scope (5A15N amplifier, 5L4N spectrum analyzer), Ballantine 320A true RMS meter, and Hewlett-Packard 5221B electronic counters and manual attenuators (350-D). Interaural phase shifts were created with a Grason-Stadler phase shifter (Model E352OB). Experiments were controlled by the microcomputer.

Procedure

The experiment was conducted in three stages. Stage 1 consisted of a BMLD measurement to ensure that the subjects participating in the pitch-recognition experiment were sensitive to the interaural phase manipulation. Stage 2 was a pitch-recognition experiment using the ABRM paradigm in combination with the BMLD procedure.

Stage 3 was a lateralization recognition experiment using a procedure similar to that used for pitch recognition.

Stage 1: Binaural masking-level difference. The BMLD for each subject was obtained using a 2IFC adaptive procedure (Levitt, 1971). The 2-down and 1-up algorithm yielded a masked threshold at 71%. Each subject's final threshold was based on the mean of three threshold estimates. Each estimate, in turn, was derived from the last six turnarounds in the run. A practice run preceded each data-collection session. A diotic 500-msec noise burst (near 15-dB spectrum level) was presented during each listening interval. The 20-msec 440-Hz signal occurred, at random, in the first or second interval and was temporally located in the center of the noise burst. The ISI and the intertrial interval (ITI) were 500 msec. The BMLD was defined as the difference in decibels between NOSO and NOS π .

Stage 2: Pitch recognition. In this experiment, we examined the effect of localization (lateralization) as a cue for pitch recognition. If stimulus features pertaining to perceptual location were made available, could a subject take advantage of the information? This possibility was examined by combining a 2IFC task and a BMLD stimulus configuration with the ABRM paradigm. The combination yielded the listening conditions necessary to answer the question.

Each interval of the 2IFC task contained the binaural stimuli required for the ABRM paradigm; namely, a tonal signal followed, after a variable delay, by the interference tone (1/7 of the trials contained no interference tone). Both the signal and interference tone were embedded in a 15-dB spectrum level diotic broadband noise (NO). When the interference tone was presented, following the signal, it was either interaurally in phase (IO) or interaurally shifted by 180° (I π). The combination of interference tone and diotic noise yielded a homophasic (NOIO) and an antiphase (NOI π) condition. The signal was always in phase with the diotic noise (NOSO) and localized at midline. The combination of these listening conditions permitted the examination of lateralization cues on pitch-recognition performance.

The ISI between signal and interference tone was selected at random prior to each trial and was 0, 20, 40, 80, 160, or 300 msec. The exception was the aforementioned control condition, where 1/7 of the trials had no interference tone following the signal. The latter condition was used to monitor the subject and maintain an unmasked performance near 80%.

The frequency of the interference tone was 440 Hz. The signal frequencies, perceived as high or low in pitch, were symmetrically spaced about the interference tone. If the subject's performance, on trials without the interference tone, varied by more than 5%, the signal frequencies were adjusted to maintain performance at 80%. Data collection began when the subject's performance on the ABRM task had stabilized to near 80%. Each set of listening conditions required eight data-collection sessions. A session consisted of five blocks, with 32 trials per block. Thus, the estimated percentage correct for each ISI was based on 160 trials. The interference tone was either antiphase or homophase throughout a block of trials. The higher pitched signal, defined as the signal with the higher frequency, occurred at random in either the first or second interval. A trial lasted 2.0 sec not counting a 1.5-sec ITI. Signal and interference tone were 65 dB SPL. Feedback was provided following each trial. Data for each subject were collected following approximately 5,000 practice trials.

In summary, each trial consisted of two intervals, the first one containing a higher pitched signal followed by the interference tone and the second one containing a lower pitched signal followed by the interference tone. In all cases, the interference tone had the same phase shift (180° or 0° degrees) throughout the block of 32 trials.

Stage 3: Lateralization recognition. In this experiment, we examined subjects' abilities at recognizing (localizing) interaural phase differences in a 2IFC ABRM paradigm. The stimulus intensities, background noise, and temporal sequencing of events in Stage 3

were the same as in Stage 2. However, whereas the phase of the interference tone was manipulated in Stage 2, the phase of the signal was varied in Stage 3. In addition, the interference tone was always homophasic.

On each trial, the subject did not judge the pitch of the signal; rather, the task was to report the interval containing the phase-shifted signal. In one interval, the signal was homophasic; in the other, it was heterophasic. The interaural phase of the heterophasic signal was adjusted, for each subject, to obtain an approximate 80% correct performance when no interference tone was present. The ISIs, 0, 20, 40, 160, or 300 msec, were randomly selected prior to each trial. On 1/6 of the trials, no interference tone was presented, to permit the monitoring of the 80% performance criterion. The frequencies of both the signal and the interference tone were 440 Hz. A session consisted of six blocks of 30 trials each. Each subject participated in eight sessions. Mean performance for each session was determined by the last five blocks. Thus, the percentage correct for each ISI value was calculated on the basis of a total of approximately 200 trials per datum.

RESULTS

Stage 1

The results of the BMLD portion of the experiment demonstrated sizable masking-level differences (MLDs) for all subjects who participated in the pitch-recognition task. The MLDs were 9.56, 9.86, and 10.27 dB for R.S., F.D., S.M., respectively. The subjects demonstrated the expected sensitivity to the phase manipulation.

Stage 2

Individual and mean data for 3 subjects, S.M., R.S., and F.D., are displayed in Figure 1. The data were analyzed via the BIOMED statistical package with the Greenhouse Geisser correction. The 6×2 repeated measures ANOVA (ISI and phase, respectively) yielded a significant main effect for ISI [$F = 13.54$, $p < .001$]. The phase main effect and the interaction of phase with ISI were not significant. In summary, the data reflect a significant increase in pitch recognition as ISI increased, regardless of the interaural phase of the interference tone.

Stage 3

In contrast with the pitch-recognition data, the lateralization portion of the experiment yielded markedly different results. Figure 2 shows individual mean performances. The BIOMED statistical analysis (repeated measures on ISI) revealed no significant results. The silent intervals between signal and interference tone had no effect on the subjects' performances. The subjects' percentages correct, averaged across ISI values, were .74, .71, .64, and .72 for D.S., S.M., L.H., and F.D., respectively. The percentages correct without the interference tone were .79, .86, .84, and .77 for the same subjects.

It is interesting that 2 subjects had to be eliminated from participation in Stage 3. They were unable to reliably discriminate the homophasic signal from the heterophasic signal in trials where no interference tone was presented. Performance for these subjects failed to reach the mandatory 80% criterion even when the interaural phase

Figure 1. Individual pitch recognition for subjects R.S., F.D., and S.M. Percentage correct is plotted as a function of interstimulus interval. Signals were homophasic and the interference tone was either homophasic or antiphasic. Signal frequencies used to obtain 80% correct were 6, 10, and 6 Hz for subjects R.S., F.D., and S.M., respectively. The X indicates performance without the interference tone. Vertical bars indicate one standard error about each point.

difference was 180° . A third subject, S.M., completed Stage 3, but only with great effort. He was able to reach the 80% criterion with an interaural phase of 150° and showed little improvement with an interaural phase of 180° . When one considers the ease with which other subjects met the 80% criterion (approximately 35° interaural phase difference), an explanation for the performance of the other 3 subjects seems obligatory. The search for such an explanation is, however, elusive. One possible factor that may be related to this failure is handedness. The 3 subjects who had problems recognizing the phase-shifted signal were ambilateral. The subjects who could perform the task with relative ease were all right-handed. Although the evidence linking recognition of a phase-shifted stimulus to handedness appears to lead to a desultory explanation, handedness and hemispheric dominance effects have been found in other auditory tasks (e.g., see Clifton & Freyman, 1989). Research along these lines warrants our interest.

Figure 2. Individual lateralization recognition for subjects D.S., S.M., L.H., and F.D. Percentage correct is plotted as a function of interstimulus interval. The interference tone was homophasic. The X indicates performance without the interference tone. Vertical bars indicate one standard error about each point.

DISCUSSION

As noted in the introduction, research on ABRM has been controversial. Data from previous studies appear to indicate, with an occasional disagreement, that a decrease or absence of ABRM could occur for at least four reasons: (1) high levels of training (Leshowitz & Cudahy, 1973; Loeb & Holding, 1975), (2) fixed or blocked ISI conditions or interference-tone frequencies (Hawkins & Presson, 1977), (3) use of the 2IFC psychophysical method (Yost, Berg, & Thomas, 1976), and (4) low stimulus uncertainty (Watson et al., 1975; Yost et al., 1976). The present study incorporated two of these four considerations; namely, highly trained subjects, and the 2IFC psychophysical procedure. Simultaneously, we randomized the ISI on each trial and produced a relatively high level of uncertainty. The latter conditions would supposedly lead to the classical ABRM effects; namely, an increasingly better performance as a function of increasing ISI. The introduction of yet another factor, the antiphase lateralization "cue," into the pitch-recognition task was expected to improve performance across all ISI

values if the processing model suggested by Massaro was incorrect. Massaro (1973) hypothesized that the PAS was a single-channel central storage where all acoustic features associated with a given stimulus may be overwritten prior to an optimal 250-msec processing interval. The results of the pitch-recognition experiment (Stage 2) demonstrated that there was interference in pitch recognition as a function of ISI. As the model suggested, subjects apparently did not use the available lateralization information to improve their performances. In both interaural phase conditions, pitch recognition was poor at short ISIs, and it improved as ISI increased.

The observation of backward recognition masking in the present study is in conflict with Yost et al. (1976), who failed to observe any masking effects using a 2IFC ABRM procedure. Because it has been hypothesized that ABRM will not occur in conditions of low stimulus uncertainty (Watson et al., 1975; Yost et al., 1976), the Yost et al. data may be explained by their use of only two ISI values in each of their experiments; thus, stimulus uncertainty was low. It is interesting that the Massaro model makes no statement regarding PAS disruption as a function of stimulus uncertainty. The model implicitly assumes that disruption should occur regardless of a subject's expectations. It should make no difference for performance if a block of trials contains one ISI or several; the information in the PAS should still be disrupted.

Although ABRM was exhibited in the pitch investigation, the interaural phase of the interference tone had no effect on performance. Lateralization cues provided by the interference tone may have gone unused for one of two possible reasons: either the lateralization features were in a separate and independent channel, or the interference tone seriously disrupted all acoustic features at short ISIs (as was postulated by the model). The data of Stage 2 do not provide a definitive answer to either of these possibilities. On the other hand, when one considers the data from Stage 3, the lateralization-recognition experiment, the viability of the model bears serious scrutiny. The latter data show a high degree of lateralization recognition, regardless of the ISI. When the feature to be recognized was lateralization rather than pitch, the task was relatively simple, regardless of the ISI. The incongruity between lateralization and pitch recognition, using the same paradigm, presents an inconsistency in the aforementioned model. According to the model, the stimulus features for both lateralization and pitch should have been degraded similarly by the interference tone, particularly at short ISIs.

Another possibility is that performance on lateralization tasks was better than performance on pitch recognition because pitch processing required memory comparisons, whereas lateralization recognition required only a gross judgment of general location. This possibility was presented by Massaro, Cohen, and Idson in 1976. Their paradigm involved differences in intensity to evoke lateralization, rather than binaural phase shifts. As in the present study, their results showed better performance on lateralization than on pitch-masking tasks. However, they

maintained that their paradigm did not provide the opportunity for a memory comparison in the lateralization task. In the present study, however, both pitch and lateralization data were obtained with the same paradigm. Thus, the difference in performance that was found between lateralization and pitch recognition cannot be accounted for by appealing to differential memory comparisons.

Of some slight concern in both experiments was the possibility that a loudness difference may exist between heterophasic and homophasic stimuli. To ensure that recognition-masking effects were not confounded by loudness cues, a post hoc experiment was conducted to determine if the antiphasic interference tone was perceptually louder than the homophasic interference tone. Interestingly, when the tonal stimuli were 65 dB SPL, there were no loudness differences for any of 4 subjects in a simple loudness-balancing task between the homophasic and antiphasic stimuli. At masked threshold, however, the MLDs for these same subjects averaged approximately 11 dB SPL. The binaural phase effects apparent at masked thresholds are dissipated at these suprathreshold levels (Townsend & Goldstein, 1972).

In conclusion, the model of pitch processing examined in this study does not adequately explain differences in performance found between lateralization and pitch-recognition experiments. Juxtaposed with this inadequacy are other data that show white noise to be a less effective interfering stimulus than a sinusoid (Kallman & Brown, 1986). According to the model, a white noise and a sinusoid should degrade performance similarly. This discrepancy lends further credence to the view that the PAS cannot be explained in terms of a simple single-channel store. The most parsimonious conclusion is that the model is inadequate as a description of auditory processing.

REFERENCES

- CLIFTON, R. K., & FREYMAN, R. L. (1989). Effect of click rate and delay on breakdown of the precedence effect. *Perception & Psychophysics*, **46**, 139-145.
- HAWKINS, H. L., & PRESSON, J. C. (1977). Masking and preperceptual selectivity in auditory recognition. In S. Dornic (Ed.), *Attention and performance VI* (pp. 195-211). Hillsdale, NJ: Erlbaum.
- KALLMAN, H. J., & BROWN, S. C. (1986). Auditory backward recognition masking with well-practiced listeners. *Journal of Auditory Research*, **26**, 239-246.
- KALLMAN, H. J., & MASSARO, D. W. (1979). Similarity effects in backward recognition masking. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 110-128.
- KALLMAN, H. J., & MASSARO, D. W. (1983). Backward masking, the suffix effect, and preperceptual storage. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 312-337.
- LESHOWITZ, B., & CUDAHY, E. (1973). Frequency discrimination in the presence of another tone. *Journal of the Acoustical Society of America*, **54**, 882-887.
- LEVITT, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, **49**, 467-477.
- LOEB, M., & HOLDING, D. H. (1975). Backward interference by tones or noise in pitch perception as a function of practice. *Perception & Psychophysics*, **18**, 205-208.
- MASSARO, D. W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology*, **85**, 411-417.
- MASSARO, D. W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, **79**, 124-145.
- MASSARO, D. W. (1973). A comparison of forward versus backward recognition masking. *Journal of Experimental Psychology*, **100**, 434-436.
- MASSARO, D. W. (1975). Backward recognition masking. *Journal of the Acoustical Society of America*, **58**, 1059-1065.
- MASSARO, D. W., COHEN, M. M., & INSON, W. L. (1976). Recognition masking of auditory lateralization and pitch judgments. *Journal of the Acoustical Society of America*, **59**, 434-441.
- TOWNSEND, T. H., & GOLDSTEIN, D. P. (1972). Suprathreshold binaural unmasking. *Journal of the Acoustical Society of America*, **51**, 621-624.
- WATSON, C. S., WROTON, H. W., KELLY, W. J., & BENBASSAT, C. A. (1975). Factors in the discrimination of tonal patterns: I. Component frequency, temporal position, and silent intervals. *Journal of the Acoustical Society of America*, **57**, 1175-1185.
- YOST, W. A., BERG, K., & THOMAS, G. B. (1976). Frequency recognition in temporal interference tasks: A comparison among four psychophysical procedures. *Perception & Psychophysics*, **20**, 353-359.

(Manuscript received March 27, 1989;
revision accepted for publication September 25, 1989.)