Backward recognition masking as a general type of interference in needed poststimulus processing

RICHARD E. PASTORE, JEREMY R. GASTON, AND MELODY S. BERENS

Binghamton University, Binghamton, New York

Auditory backward recognition masking (ABRM) has been argued to reflect interference in the storage and/or processing of a short-lived sensory form of information and has been viewed as a relatively invariant attribute of auditory pitch processing for very brief stimuli that are minimally separated in frequency (ΔF). In contrast, the present study demonstrates that ABRM reflects interference with several basic principles of auditory processing. Measured in terms of target tone duration, rather than ΔF , ABRM is demonstrated for target stimuli representing the interval of a musical fifth and masker–target stimulus intervals of a musical third, with thresholds ranging from approximately 22 to 55 msec and psychometric functions that are indicative of more than one contributing factor. On the basis of common underlying principles, the possibility that the threshold for the identification of temporal order of onset reflects ABRM and possible implications for the perception of complex stimuli, including speech, are discussed.

Auditory backward recognition masking (ABRM) is the interference in the recognition of some sound attribute due to the presentation of a subsequent sound. In a typical ABRM paradigm, illustrated in the upper panel of Figure 1, the listener is asked to identify the pitch of alternative brief (10- or 20-msec) stimuli that, when presented in isolation, are close to the difference limen (DL) for recognition. When the target stimulus is followed by a fixed-property contextual stimulus, recognition accuracy is reduced by an amount inversely related to masker delay. Investigated over the last decade largely in applied contexts as a fixed characteristic with neurological correlates (e.g., McArthur & Bishop, 2002) or individual differences that may be correlated with specific cognitive attributes—for example, reading disabilities (Griffiths, Hill, Bailey, & Snowling, 2003; McArthur & Hogben, 2001) or intelligence (Deary, 1995)—ABRM would appear to be of relatively limited relevance to the perception of complex stimuli. In fact, the gating paradigm (e.g., Grosjean, 1996), commonly used in the speech literature to evaluate when auditory information is first perceptually available, makes the implicit assumption that later occurring portions of the stimulus are irrelevant and, thus, assumes an absence of any ABRM. Rather than being a relatively minor form of specific, relatively invariant interference with limited general relevance, we believe that ABRM reflects the interaction of several general principles of auditory information processing that are common to simple and complex listening situations. The present study is an initial attempt to develop the broader generality and importance of the principles underlying this phenomenon.

Auditory Backward Recognition Masking

ABRM was originally investigated from two different research perspectives. The goal for Massaro (1970) was to evaluate the duration of the short-term auditory sensory store, with the notion that the contextual stimulus disrupted the sensory information still being processed in the pitch identification task. The goal for Ronken (1971, 1972) was to investigate the frequency DL of brief 10-msec tones. These and the other past ABRM studies always investigated brief target stimuli that were very similar in frequency, measuring either the frequency DL (thus, constant accuracy) or accuracy for a constant (near the DL) frequency difference. For relatively naive listeners, accuracy is significantly reduced at relatively short masker delays (e.g., interstimulus interval [ISI] < 50 msec), systematically improving to an asymptote that often still indicates some interference at ISI values of 250–350 msec (e.g., Massaro, Cohen, & Idson, 1976). When the frequency DL is measured, there is a relatively small change in the DL, and interference is found only up to ISI values of 100–150 msec (e.g., Foyle & Watson, 1984; Leshowitz & Cudahy, 1973). These differences in magnitude of effect were the subject of debate about the more appropriate psychophysical measure of ABRM and the theoretical importance of the phenomenon (Foyle & Watson, 1984; Leshowitz & Cudahy, 1973). The importance of the phenomenon and what is conveyed by the different measures can be understood in terms of the roles played by the principles and sequence of stages involved in basic auditory information processing.

The underlying principles are best understood with a perspective on the complex nature of ABRM. The phe-

R. E. Pastore, pastore@binghamton.edu

Figure 1. Conceptual diagram of relevant stimulus paradigms. The upper panel summarizes stimuli in a typical auditory backward recognition masking paradigm. The middle panel summarizes the stimuli for the present study. The lower panel summarizes stimuli for a temporal order judgment paradigm. DL, difference limen; ISI, interstimulus interval; SOA, stimulus onset asynchrony.

nomenon is distinct from backward (detection) masking in that target tones in ABRM are always detectable and there are major differences in both the time course and dependence on masker properties (e.g., Bland & Perrott, 1978; Elliott, 1962; Moore, 2003). The nature of masker and target tone properties provides some indication of the possible locus (or multiple loci) of the interference contributing to ABRM. Energy in a peripheral filter (critical band) that is the same as that for the target is not necessary; although ABRM has been found for maskers intermediate

in frequency to the target tones (as depicted in the upper panel of Figure 1), it has been either minimal (Holding, Loeb, & Yoder, 1972) or has not been found (e.g., Sparks, 1976) for a noise masker. ABRM has also been found for maskers that differ somewhat in frequency from the target tones, with magnitude of interference sometimes found to be greater than (e.g., Hawkins & Presson, 1977; Massaro et al., 1976) and sometimes less than (Sparks, 1976) that found with target-centered maskers. In addition, ABRM has been found for monaural target tones with contralateral masker (e.g., Hawkins & Presson, 1977). Another indication that ABRM probably reflects the interaction of several underlying factors is that the magnitude of ABRM varies as a function of listener skills, with musically trained listeners, and all listeners with extensive practice, usually exhibiting significant reduced magnitudes, but not the complete elimination, of ABRM (Loeb & Holding, 1975; Sparks, 1976).

Principles Underlying ABRM

There are several common, related working principles underlying ABRM. First, ABRM can occur only if there is processing after the stimulus offset. More generally, auditory information is not processed instantaneously but, rather, is delayed in time. Furthermore, simple temporal proximity to the target stimulus is not sufficient, since no interference is found when the masking stimulus is presented before the target stimulus. In fact, although accuracy is not altered for 20-msec target stimuli (e.g., Leshowitz & Cudahy, 1973; Massaro et al., 1976; Sparks, 1976), improved accuracy or forward recognition contrast can be found for 10-msec target stimuli (Leshowitz & Cudahy, 1973; Ronken, 1972). The nature of delayed processing is probably best represented as a series of overlapping stages that differ in the characteristics of information (e.g., Cowan, 1988; Massaro, 1977; Näätänen & Winkler, 1999).

Second, the amount of poststimulus processing needed depends on the difficulty of the perceptual task facing the listener. Task difficulty depends on a number of factors, including the quality or precision of the stimulus information, the similarity of alternative perceptual decisions, the relevant skills of the listener, and the amount of processing that can occur while the stimulus is present. Like the stimuli in the early studies of visual short-term storage (e.g., Averbach & Coriell, 1961; Sperling, 1960), ABRM target stimuli are very brief, forcing significant dependence on poststimulus processing. Because both task difficulty and efficiency of processing, both during the presence and after the offset of the physical stimulus, depend on the relevant skills of a participant, the magnitude and temporal course of ABRM should, and do, differ with musical experience and practice (e.g., Sparks, 1976).

Third, the spectral properties of auditory stimuli, such as bandwidth, develop and, thus, change as a function of time. Specifically, the spectrum of a brief segment of a sinusoid is described by a sinc, or $\sin(\omega t)/\omega t$, function, with effective bandwidth inversely proportional to duration (Cooper & McGillem, 1967). As a result, the pitch of a segment of a sinusoid begins to emerge only after a minimum duration or, for low frequencies, a minimum

number of cycles, with a very brief sinusoid heard as a click and perception changing to a tone pip and, finally, to a tone as duration is increased (Doughty & Garner, 1947, 1948; Patterson, Peters, & Milroy, 1983). In fact, the ability to identify a rough octave is limited for brief segments of tones (Robinson & Patterson, 1995). Consequently, in past ABRM studies, the difficulty for the pitch recognition processes and, thus, the dependence on poststimulus processing were due to both limited frequency separation and duration. If task difficulty is a major factor in finding ABRM, either of these factors should create the conditions needed for this phenomenon, which means that ABRM should be found when masking is mapped in terms of duration for target tones that are well beyond the frequency DL; this will be demonstrated in the present study.

ABRM As a Function of Duration for Spectrally Separated Stimuli

The only ABRM study in which target stimulus duration has been manipulated was for spectrally and perceptually very similar target and masker stimuli and, thus, was very different from the present study. Massaro (1972) measured ABRM in terms of change in d' as a function of both target stimulus duration and masker delay for a sine versus sawtooth target stimulus waveform and a square wave masker, all with a fundamental frequency (F_0) of 800 Hz. The masker and two target stimuli differed solely in presence and distribution of energy at the harmonics (e.g., Rossing, 1990). In contrast to the present study's focus on pitch judgment for large differences in frequency across the target and masker stimuli, all the stimuli in the Massaro (1972) study were identical in subjective pitch and differed solely in subtle aspects of spectral timbre.

If the interaction of the principles summarized above plays a meaningful role for a broad spectrum of auditory perceptual tasks, rather than for only very brief stimuli contrasted by little more than the frequency DL, ABRM needs to be demonstrated for stimuli that are both longer than 10–20 msec and physically more distinct. Furthermore, although ABRM has been conjectured to reflect more than low-level sensory processing (e.g., Hawkins & Presson, 1977; Loeb & Holding, 1975), additional supporting evidence would be provided by demonstrating ABRM for sets of target stimuli and maskers that are all distributed across separate critical bands. Finally, although stimulus differences of 10 or 20 msec can be important in the perception of speech (e.g., voicing contrast along a voice onset time continuum), it is common to find components of speech and other natural stimuli to be of longer duration; if ABRM types of interference are relevant to these stimuli, ABRM would need sometimes to occur in stimuli that are longer than 10 or 20 msec. The following experiments were designed to address these questions.

EXPERIMENT 1

With target stimuli separated by the musical interval of a fifth (702 cents), the basic pitch recognition task should be limited by stimulus duration. To further ensure that we were dealing with a more central, rather than a

peripheral, form of interference, the frequencies of the masking stimuli differed from each other and from the target tones by more than the estimated width of the conceptual peripheral filter (equal rectangular band, or ERB; Moore, 2003). On the basis of the summarized principles, the magnitude of dependence on poststimulus processing and, thus, susceptibility to ABRM interference should be inversely related to stimulus duration, reflecting at least two factors that change with stimulus duration: narrowing effective stimulus bandwidth (decreasing task difficulty) and opportunity for greater processing while a stimulus is physically present.

An adaptive procedure provides an efficient way to estimate thresholds (e.g., Macmillan & Creelman, 2005) and now is often used in studies of masking (see, e.g., Lutfi, Kistler, Callahan, & Wightman, 2003). In the ABRM literature, Foyle and Watson (1984) employed an adaptive procedure (Levitt, 1971) to measure the frequency DL for 20-msec target stimuli centered on 800 Hz as a function of target–masker delay intervals (ISIs) for an 800-Hz masker. Experiment 1 also used the Levitt adaptive procedure to measure target duration threshold.

In the absence of any masking stimulus, participants with reasonable listening skills who understand the task should yield thresholds that correspond to the established 8- to 10-msec threshold for pitch identification (Doughty & Garner, 1947, 1948). Significantly longer thresholds in the presence of subsequent contextual stimuli demonstrate ABRM interference. Our expectation is that for many masking stimuli, the threshold duration will approximate the longer 20-msec stimuli used in many ABRM studies but that the threshold for some maskers might be significantly greater, indicating greater generality for the basic principles that motivated this study.

Method

Participants. The participants in all three experiments were Binghamton University students who received credit in partial fulfillment of the research requirement for undergraduate classes. All were of typical college age and reported normal hearing. Each participant was run in a commercial sound chamber. For Experiment 1, data were collected for a total of 45 participants; because 4 participants failed to meet the *a priori* criterion for inclusion in the experiment, reported data are from 41 participants.

Stimuli. The middle panel of Figure 1 summarizes the stimuli. All the stimuli (targets and maskers) were synthesized at a 12-bit, 15K sample rate with onsets and offsets at waveform zero-crossings. Target tone duration (cycles) and ISI were controlled online by the computer program that ran the experiment. All the stimuli were low-pass filtered at 2 kHz and were presented binaurally through TDH-49 headphones. One-hundred-millisecond versions of all the stimuli were calibrated to be 65 dB(C) and, thus, were clearly detectable. The target tones were 250 and 375 Hz, which, when of sufficient duration, can be easily distinguished.

The eight conditions were defined by the nature of the masker. In the no-masker condition, only the target tones were presented, and the threshold for this condition served as a baseline. The remaining conditions differed solely in the nature of the masking stimulus that was present on every trial with a fixed ISI of 5 msec. The amount of masking in each of the seven masker conditions was measured as the difference in threshold between the masking condition and the no-masker baseline. Although energetic masking was either minimal or absent in prior ABRM studies, because those studies were based on very different dependent measures, we included a white noise

masker condition. The remaining masking stimuli were either single tones, designated as L (low = 200 Hz), M (middle = 313 Hz), and H (high = 469 Hz), or the paired mixture of these tone maskers—thus, L–M (low–middle), L–H (low–high), and M–H (middle–high). The relationship between each single tone masker and each adjacent target tone approximated a musical third (386 cents) and, thus, exceeded the limits of the ERB centered on the target tone. When of sufficient duration, the five tones (two target and three masking tones) should be easily differentiated and should not significantly interact at peripheral levels of the auditory system. To avoid clipping, the three mixed masker stimuli were created by digitally attenuating each of the three masker tones by 6 dB (half voltage), and then summing the digital samples.

Procedure. Each listener participated in three blocks of trials, which together took approximately 1 h to complete. All the listeners first participated in the no-masker condition, which provided a no-masking threshold baseline, experience in identifying the two target tones, and an *a priori* basis for screening out listeners who lacked either sufficient motivation or the very basic listening skills needed for participation. Each listener then participated in two of the eight possible conditions, with the pairing and ordering of conditions roughly counterbalanced across listeners (i.e., only 41 of the 56 possible ordered samples were run). There thus were data from 41 listeners for the initial no-masker condition and a minimum of 9 listeners for each of the other eight conditions.

The basic task for the initial no-masker condition was described to the participant. The participant was instructed to indicate (using two vertically arranged buttons) whether the high or the low target tone had been presented at the beginning of the stimulus. After the participant had completed the no-masker condition, each of the next conditions was described. Each task was the same, but the initial tone in the masking conditions was followed by a stimulus that did not change across trials, with listeners instructed to try to ignore the second sound. Feedback indicating the correct response (the word *high* or *low*) was provided at the end of each trial.

The duration threshold for differentiating between the two possible target tones in the context of the various subsequent (trailing) stimuli was determined using a version of Levitt's (1971) up–down procedure with a 4:2 decision rule, which estimates threshold at an accuracy rate of .686. The initial target tone duration was 100 msec. Duration then was adjusted in the following manner. When four correct responses occurred before two incorrect responses, the target was shortened; when two incorrect responses occurred prior to four correct responses, the target was lengthened. If no response was made, the trial was ignored. Target stimulus property (number of cycles) was recorded for every reversal (change of direction). The duration threshold was estimated in separate adaptive sequences for the high and the low target tones; when the computer randomly selected the target tone for the given trial, it imposed the current properties of the adaptive sequence for that target tone.

Target tone duration and changes in duration were defined in terms of complete cycles of the target tones, thus also maintaining onsets and offsets at positive waveform zero-crossings. The initial change in duration was 10 cycles of the target tone and, therefore, 40 msec for the lower tone or 26.7 msec for the higher tone. At each reversal, step size was reduced to 5, 3, and 2 cycles and, finally, became constant at 1 cycle. In addition, the minimum target duration was 1 cycle that, when reached, was treated as a reversal. The condition ended when a minimum of 15 reversals was recorded for both the high and the low target tone sequences, sometimes resulting in more than 15 reversals for one of the sequences. The first 5 reversals were treated as both practice and the initial approximation of the threshold region, with separate thresholds defined as the average number of target tone cycles across the last 10 or more reversals. Since differences across sequences reflect, in part, response bias, the separate thresholds were averaged. However, although pitch identification threshold for low-frequency tones is best reflected by number of cycles (e.g., Patterson et al., 1983), a major goal for the experiment was to make comparisons with the stimulus durations used in

the ABRM literature. To accomplish this, the threshold results were converted to duration by multiplying by the average of the period of the two target tones (4.0 and 2.67 msec, or 3.33).¹

Results and Discussion

An outlier analysis of the data for the no-masker condition was based on the criterion of three standard deviations from the average group threshold, resulting in the removal of 4 participants. With multiple conditions to compare in our mixed design and, thus, several statistical tests, we used an alpha level of .01 as the criterion for rejecting the null hypothesis. Figure 2 summarizes the average threshold (with standard error) for each of the conditions. For the initial no-masker condition (left bar in Figure 2), the average threshold was 8.4 msec (range: 8.0–12.2 msec), corresponding to 2.1 cycles of the lower target tone or 3.2 cycles of the higher target tone, which is consistent with the minimum duration needed for pitch quality to begin to emerge (e.g., Doughty & Garner, 1947, 1948). A subgroup of participants took part in this no-masker condition a second time (second bar from the left in Figure 2), with the essentially unchanged 8.1-msec threshold, indicating no effect of practice or fatigue.

All of the remaining conditions had a subsequent masking stimulus. In Figure 2, the results for these conditions are displayed in a manner that allows a direct within-group evaluation of the amount of masking. The dark portion of each bar indicates the average initial no-masker threshold for the group of participants who took part in the specific masking condition, with very little variability across groups in these baseline thresholds. For each condition, the overall height of the bar indicates the average masked threshold, and the lighter portion of each bar represents the change in threshold due to the presence of the specified masking stimulus and, thus, the magnitude of ABRM. In prior ABRM studies, a broadband noise masker produced little (Holding et al., 1972) or no (Sparks, 1976) masking. The present 11.6-msec threshold does differ significantly from the no-masker threshold [$t(11) = 4.71, p < .001$], but this difference of approximately one average target tone cycle is probably not behaviorally meaningful.

In Figure 2, the single-masker and mixed-masker conditions are separately grouped within vertical lines. All of these remaining masking conditions were expected to, and clearly did, exhibit ABRM. Among the three single-tone masker stimulus conditions, the recognition thresholds did not differ significantly from each other [$F(2,29) = 0.106, p = .36$] but, as a group, differed significantly from the no-masker thresholds [$t(31) = 5.33, p < .001$]. Thus, ABRM was clearly demonstrated for single-tone maskers, using target tone duration as the dependent measure for target tones that differed by a musical fifth. Furthermore, all of these three single-tone masker conditions' thresholds were at least somewhat longer than the longest 20-msec target tone duration used in past ABRM studies.

One goal for this study was to demonstrate that the underlying principles are relevant to complex listening conditions and, thus, beyond the interaction among brief tones. The mixed-masker conditions, grouped on the right side of Figure 2, represent a step in that direction, evaluating whether

Figure 2. Experiment 1 average threshold (and standard error) results as a function of condition. The left pair of bars is the original and replication threshold for the no-masker condition. The middle and right sets of three bars are for the single- and mixed-masker conditions, with the overall height indicating masked thresholds for each given condition. The dark portion of each bar indicates the average no-masker threshold for the subset of participants who ran that masking condition; the lighter portion of the bar thus indicates amount of masking. L, low; M, middle; H, high.

interference from somewhat more complex maskers is less than, equal to, or greater than that of the individual masker components. Within the mixed-masker conditions, there was a significant main effect of masker type [$F(2,29) = 5.91$, $p < .01$], meaning that ABRM is not a fixed property. Post hoc comparisons indicate that the L-M and L-H masker conditions did not differ from each other [$t(17) = 0.433$, $p = .67$], and the combined data from these conditions were significantly greater (longer) than the threshold for the M-H masker condition [$t(21) = 4.07$, $p < .006$]. If we make the simplifying assumption that the amount of masking is the difference in threshold caused by the presence of the masker, mixed-masker interference can be greater than (L-M masker) or equal to (L-H masker) the sum of interference produced by its components, but also (L-M masker) can be less than, but not significantly different from, either of its component tones [$t(25) = 1.29$, $p = .21$]. We thus have an indication of possible complex masker interactions that will be addressed further in terms of the psychometric functions evaluated in Experiments 2 and 3.

EXPERIMENT 2

Like many modern studies of masking (e.g., Kidd, Mason, Deliwala, Woods, & Colburn, 1994; Neff & Dethlefs, 1995; Oh & Lutfi, 1998), Experiment 1 used an adaptive procedure to directly estimate thresholds corresponding to $p = .69$. Adaptive procedures efficiently provide direct estimates of thresholds and, thus, the magnitude of masking in physical terms that are easily understood. However, studies in which ABRM has been quantified in terms of reduction in accuracy typically have indicated considerably broader, and possibly greater, effects than those indicated by thresholds. These differences prob-

ably can be attributed to many different factors, including psychophysical procedure (e.g., Yost, Berg, & Thomas, 1976) and participant experience (e.g., Loeb & Holding, 1975; Sparks, 1976). Which measurement approach is more appropriate? Although this question seems simple, the answer is not. When psychophysics is used to study simple sensory-based relationships, such as critical bandwidth from masking (e.g., Fletcher, 1940), the psychometric functions across conditions are similar approximations of the same underlying phi-gamma (cumulative normal) distribution shifted along the stimulus axis. With functions that differ only in location along the magnitude axis, the difference in thresholds reflects the amount of displacement independently of the accuracy criterion that defines threshold. Similar psychometric functions would be expected if ABRM reflects a fixed property of auditory processing. However, the psychometric functions for ABRM do not approximate a phi-gamma distribution (e.g., Foyle & Watson, 1984) and, thus, probably reflect the interaction of a number of different types of auditory information processing. Furthermore, the estimated magnitude of interference will depend on the criterion for threshold (i.e., amount of masking will differ for a .69 vs. a .77 or .84 criterion—or $d' = 1.0$, 1.5, or 2.0—for threshold).

A similar concern has been expressed over the best approach for investigating (simultaneous) informational masking that also probably reflects multifaceted, higher level processes. Although often quantified by thresholds estimated using adaptive procedures, there are concerns that a single measure may not fully convey the broader pattern of interference that might be reflected in psychometric functions (e.g., Durlach et al., 2005; Lutfi et al., 2003). This does not necessarily mean that quantitative estimates of masking based on thresholds are incorrect. Any

phenomenon can be best understood by a multifaceted approach that reflects both direct estimates of thresholds and generation of psychometric functions.

Experiment 2 evaluated ABRM psychometric functions for all of the major conditions in Experiment 1, except for the noise masker. Although the single-tone maskers in Experiment 1 did not differ in estimated magnitude of effect, the pairs of maskers did, making it quite possible that the nature and level of interference differed across the single-tone maskers. It thus would not be surprising to find somewhat different psychometric functions across the single-tone maskers. In addition, although the single maskers were components in the mixed-masker stimuli, the psychometric functions for the mixed maskers depended on the way they separately and together altered the pitch processing and decision strategies. Thus, the single- and mixed-masker condition psychometric functions may have differed in ways not reflected in the Experiment 1 threshold results. Although essentially a replication of the conditions in Experiment 1, there were several modifications: The no-masker condition was evaluated for each participant only at the beginning of the session, and masker type was now strictly a between-subjects variable.

Method

Participants. A total of 190 participants were run to generate data for 24 participants in each of the six masker conditions, or a final total of 144 participants.

Stimuli. A new set of stimuli with the same frequency properties as those in Experiment 1 were synthesized (16-bit, 44-kHz sampling rate) using Adobe Audition 2.0, again edited to begin and end at a waveform zero-crossing. The 250- and 375-Hz target tones were each created with durations from 12 to 96 msec in 12-msec steps, allowing both tones to begin and end at waveform zero-crossings. The shortest three durations were used in the no-masker condition. For each masker condition, the digital code for each of the target tones was combined with the ISI (5 msec of silence) and the masker to create a set of 16 stimuli, with masker duration now 50 msec. The stimuli were presented binaurally over Sennheiser HD 260 headphones at approximately 65 dB(C).

Procedure. Participant instructions, stimulus selection and presentation, data collection, and all other aspects of the task were programmed using E-Prime 1.0 (Psychology Software Tools, 2002). Each participant ran two (blocked) conditions: first, the no-masker baseline; then one of the six masker conditions, with each of these conditions preceded by a brief 12-trial practice block that consisted of three presentations of each of the two target tones at the two longest target durations (24 and 36 msec for the no-masker condition and 84 and 96 msec for the masking condition). For the no-masker condition, the participants were told that they would be presented with one of two very brief tones, with their task being to indicate whether the tone was the higher or lower in pitch, responding with keys marked "H" and "L." For the masking condition, they were told that the same brief tones would be followed by a longer stimulus that would never change; their task again was to identify the initial tone as being the high or low pitch and to ignore the final tone. For both the no-masker and the masking conditions, every possible stimulus was presented 20 times, with the program determining the random order of presentation. Feedback indicating the correct response (high or low target) was provided at the end of every trial. In Experiments 2 and 3, the response, feedback, and intertrial intervals combined lasted 5 sec.

Results and Discussion

The goals for the no-masker condition were to provide the participants with experience in judging the relative target tone pitch and to screen out participants who had

difficulty making the easiest (36-msec duration) of the expected suprathreshold pitch judgments in the absence of a masker. In Experiments 2 and 3, if a participant's performance for these no-masker condition stimuli was below 75% correct, that participant was eliminated from the study, with a new participant tested.

For each masker condition, accuracy at each target tone duration was averaged across the 24 participants to generate the average psychometric function. Because the no-masker condition did not vary across masker conditions, the no-masker psychometric functions should reflect similar baseline performance. The left panel in Figure 3 shows these psychometric functions, and a one-way ANOVA indicated no significant effect of condition [$F(5,426) = 1.8, p = .11$], indicating that the groups did not differ in relevant pitch identification abilities.

The right panel in Figure 3 shows the average psychometric functions for the single- and complex-masker conditions, with the no-masker condition results pooled across all 144 participants plotted as a baseline reference. If one looks only at the data for the 12-msec target tone duration, where accuracy in all the conditions was below 75%, there would seem to be little ABRM in the two masker conditions (M and M-H) that exhibited the least ABRM in Experiment 1 (where threshold criterion was 0.69) and moderate to considerable ABRM in the remaining masker conditions. However, the full psychometric functions convey a somewhat different picture.

To evaluate the presence of ABRM in the masking conditions, a mixed design ANOVA with two within-subjects factors (masker presence/absence and duration) and one between-subjects factor (masker type) was run on the data from the three shortest target tone durations (where data exist for all the conditions). Because the sphericity assumption always was violated for the factor of target duration in the psychometric functions for Experiments 2 and 3, a Greenhouse-Geisser correction was used for all within-subjects effects. A target duration effect simply reflects an increasing psychometric function and was always highly significant ($p < .0001$); the inferential statistics for target duration thus will not be reported. As was expected, there was a significant main effect of masker presence [$F(1,138) = 264.12, p < .0001$], indicating reduced accuracy in the ABRM conditions. The significant main effect of masker type [$F(5,138) = 8.1, p < .0001$], coupled with the lack of a significant group effect for the no-masker condition, indicates differential ABRM across the maskers. On the basis of the psychometric functions, it is not surprising that there was also a significant two-way interaction between duration and masker presence [$F(8,224) = 4.3, p < .0001$], as well as between masker presence and masker type [$F(5,138) = 14.8, p < .0001$], and finally a significant three-way interaction between duration, masker presence, and masker type [$F(9,240) = 4.1, p < .0001$].

Differences among ABRM conditions were evaluated with a one-way ANOVA using data at all target tone durations. As was expected, there was a significant effect of ABRM condition [$F(5,1146) = 48.63, p < .0001$]. Tukey's HSD post hoc tests revealed no significant differ-

Figure 3. Experiment 2 no-masker (left panel) and masker (right panel) average psychometric functions for separate groups of participants run in each masking condition. The overall average no-masker psychometric function is plotted with the masking conditions (right panel) as a baseline reference. L, low; M, middle; H, high.

ence across the four psychometric functions that exhibited the lower amounts of ABRM: the single L and M masker and the mixed L–M and M–H masker (collectively, the small ABRM) conditions, with the largest difference between conditions not greater than $M_{\text{diff}} = .041$ ($p = .90$). Although accuracy in these conditions at the 96-msec target tone duration was somewhat below no-masker condition accuracy at the 36-msec target tone duration, the difference was not significant for two of these conditions [L, $t(23) = 2.66$, $p = .007$; M, $t(23) = 0.78$, $p = .22$; L–M, $t(23) = 1.09$; $p = .14$; M–H, $t(23) = 4.29$, $p < .0001$]. In contrast, the H masker condition differed significantly from the small ABRM conditions, with the smallest difference across these groups not less than $M_{\text{diff}} = .11$ ($p < .001$). The comparison between the L–H masker condition and the small ABRM conditions also was significant, with the smallest difference not less than $M_{\text{diff}} = .146$ ($p < .001$). Finally, although it would appear in Figure 3 (right panel) that accuracy was lower at most durations for the L–H masker than for the H masker, the difference was not significant ($M_{\text{diff}} = .039$, $p = .26$). Thus, Experiment 2 indicated significant ABRM for all maskers, with significantly greater ABRM for the H and the L–H maskers.

In a broad sense, most of the Experiment 1 thresholds (Figure 2) are similar to the Experiment 2 psychometric functions (Figure 3, right panel) at $p = .69$, but the overall psychometric functions for some conditions do indicate differences not reflected in the thresholds. Although the threshold for the H masker was similar to those for the other single-tone maskers, the full psychometric function was always well below those for the other single-tone and two of the mixed-masker conditions. In Experiment 1, threshold was highest in the L–M masker condition, but the psychometric function indicates that, although there was a significant reduction in accuracy for the shortest target tone duration, accuracy was no different from that for most of the maskers at longer tone durations. Finally,

there are two noteworthy properties of the psychometric functions. The Experiment 1 thresholds had indicated that ABRM usually increased threshold stimulus duration for identifying pitch to 20–25 msec for most maskers, but to longer durations for at least some maskers, and the steeper portions of the psychometric functions are generally consistent with these estimates. However, all the psychometric functions exhibit long upper tails with somewhat depressed accuracy; for the 96-msec target duration, accuracy ranged from somewhat to greatly below that achieved for the 36-msec target duration in the no-masker condition. Furthermore, the functions for the H and L–H maskers also are markedly flat across the broad range of target tone durations investigated. Across conditions, these results imply that ABRM probably reflects more than one contributing factor. There clearly is a large interference effect for all maskers when target duration is brief. For all maskers, there also is a form of interference even with target tones not only long enough to be relatively narrow-band, but also to allow significant processing while physically present. The latter probably indicates interference at a later or more central stage of processing (e.g., Hawkins & Presson, 1977), even possibly at the decision stage.

EXPERIMENT 3 ABRM As a Function of ISI

Independently of target tone properties, masker properties, and psychophysical procedure, when accuracy is mapped across ISI, the resulting function typically deviates markedly from a phi-gamma distribution. Furthermore, not only are the psychometric functions used to estimate ABRM in terms of ΔF asymmetric, their slopes also differ as a function of ISI (e.g., Foyle & Watson, 1984). What aspects of interference might account for such patterns of results? Accuracy in judging the pitch of the target stimuli probably reflects at least three different factors:

basic task demands, the availability of poststimulus processing, and some aspect of the relationship between the target and contextual or masking stimuli. In past ABRM studies, basic task difficulty has been determined by both the short duration and the high spectral similarity of the target tones. Mapping accuracy or the frequency DL as a function of ISI was essentially done to investigate the effectiveness of poststimulus processes in compensating for the task difficulty. In the present study, basic task demands were determined by target tone duration, but not by high spectral similarity, and ISI has, so far, been constant. In principle, an equivalent type of compensation would exist for the conditions in the present study, but probably with some important differences. Because the stimuli were not similar in frequency and stimulus duration was manipulated, the ability to distinguish among the target tones and the masker may have involved later or higher stages of processing and, thus, interference, which would be reflected as a general reduction in accuracy over a broader range of masker delays. As a result, evaluating the ABRM psychometric function across ISI for one of the significant ABRM conditions could provide insights into the factors contributing to ABRM interference. In Experiment 3, psychometric functions (accuracy as a function of target tone duration) were generated as a function of masker delay for the L–H mixed masker, which had produced significant interference over a broad range of target tone durations.

On the basis of the no-masker psychometric functions in Experiment 2, the target tone durations for this condition again varied in 12-msec steps, but to a longer, 60-msec maximum. Since the Experiment 2 psychometric function for the L–H masker condition was shallow across the range of studied target tone durations, that range was expanded by using 24-msec, rather than 12-msec, steps up to a maximum of 120 msec. The no-masker and masker conditions were mixed within a single block of trials.

Method

Participants. Although it is not unusual to find reduced accuracy on masker-absent trials when they are mixed with masker-present trials, rather than run in separate blocks (Massaro et al., 1976), inclusion of a participant in the experiment was based on a no-masker-condition a priori accuracy criterion that was the same as that in Experiment 2. On this basis, 15 listeners were removed from the experiment, with data reported for 17 listeners in each of the three ISI conditions.

Stimuli. Except for ISI and the distribution of target tone durations, the stimuli, again created using Adobe Audition 2.0, were identical to those in the L–H mixed condition in Experiment 2. A total of 40 stimuli were created, 10 for each of the four conditions. The no-masker-condition stimuli consisted of 12-, 24-, 36-, 48-, and 60-msec segments of each of the two target tones. In the masking conditions, there were five target tone durations—24, 48, 72, 96, and 120 msec—with each target tone digitally combined with the masking stimulus and an intervening silent (ISI) interval of 5, 25, or 50 msec. There thus were three sets of masking condition stimuli that differed only in ISI.

Procedure. As in Experiment 2, each listener generated data for the no-masker condition, plus one of the three masker ISI conditions. The listener initially was told that the task always would be to identify the relative pitch of a brief target tone presented either alone or as the first of two tones presented on each trial, with a diagram illustrating the stimuli. After familiarization with the sound chamber, all communication with the listener, including instructions for each of three blocks of trials and trial-by-trial feedback, was handled

automatically by the E-Prime program through a monitor situated in front of the participant.

Each participant was run in two brief initial blocks of 12 randomly ordered training trials. In the first block, the participant passively listened to three repetitions of the 48- and 60-msec target tones, with relative pitch (either *high* or *low*) displayed on the computer monitor. In the second training block, the participant was presented with two repetitions of the 36-, 48-, and 60-msec target tones but now was instructed to identify the relative pitch of the tones, with feedback (*correct* or *incorrect*) provided at the end of each trial. All responses were made using keys labeled “H” and “L.” Data then were collected for 10 repetitions of the 10 target stimuli in both the no-masker and the masker (specific ISI) conditions, with stimulus order randomly determined by the computer program. Because of the mixed block design, the listener was told that, on some trials, the target tone would be followed by a second sound but that they were to try to ignore the second sound and indicate only whether the first sound was the high or the low tone.

Results and Discussion

The average no-masker condition psychometric functions are shown in the left panel of Figure 4 and indicate that asymptotic performance was reached at the 48-msec target tone duration. These data were analyzed using a repeated measures ANOVA with one within-subjects factor (target duration) and one between-subjects factor (masker ISI condition). Because the no-masker condition stimuli were identical across the groups that ran the various masker conditions, the results should be equivalent, and this is precisely what we found, with no main effect of condition [$F(2,48) = 1.38, p = .26$] and no interaction [$F(4,107) = 1.26, p = .28$].

The no-masker condition data were averaged across the 51 participants and are plotted as a baseline psychometric function in the right panel of Figure 4, along with the psychometric functions from the three masker conditions. A repeated measures ANOVA with one within-subjects factor (target tone duration) and one between-subjects factor (masker ISI) was performed on the masked data, indicating a significant main effect of masker ISI [$F(2,48) = 4.22, p = .021$] and a significant interaction [$F(7,159) = 2.5, p = .02$].

Relative to the no-masker condition, all of the masker ISI psychometric functions, at a minimum, exhibit a shift toward longer target tone durations. Similar to the L–H masker function in Experiment 2, the function with the same 5-msec ISI is relatively flat and depressed across the full range of target tone durations, and even at the longer target durations now evaluated, pitch identification is still depressed, relative to baseline asymptote. Across the three values of ISI evaluated, there was an apparent dissociation between long- and short-duration target tones as a function of ISI that is suggestive of there being at least two different interference components. With the 25-msec ISI, the accuracy for short target tones remained depressed, but accuracy was improved at the longer target tone durations. Increasing ISI to 50 msec resulted in little change in the long tail, but now accuracy for the briefer target tones was improved. Using the no-masker 60-msec target tone duration as a baseline, accuracy at the longest 120-msec target tone duration does differ significantly for the 25-msec ISI condition [$t(16) = 2.24, p = .02$], but not quite for the 50-msec ISI condition [$t(16) = 1.87, p = .04$; two post hoc tests require $\alpha = .025$].

Figure 4. Experiment 3 no-masker (left panel) and masker (right panel) average psychometric functions for separate groups of participants run in each interstimulus interval (ISI) masking condition, with the overall average no-masker function again plotted as a reference for the masking conditions.

GENERAL DISCUSSION

Nature of ABRM

The original goal of Massaro's (1970) ABRM study, to evaluate the duration of auditory short-term sensory storage, required the implicit assumption that pitch recognition processing must occur after the offset of the physical stimulus. The shift to mapping the interference function in terms of threshold for (fixed accuracy) target differences (e.g., Leshowitz & Cudahy, 1973) added an implicit assumption of a trade-off between target properties and processing time; a greater difference in target properties (i.e., ΔF or, in the present study, target duration) reduced task demands and, thus, the amount of needed poststimulus processing. The theoretical focuses then became which approach best quantified ABRM, the nature of the specific underlying cause, and the perceptual importance of a phenomenon limited to highly similar (near frequency DL) stimuli (e.g., Foyle & Watson, 1984; Hawkins & Presson, 1977; Leshowitz & Cudahy, 1973; Massaro, 1977). Although parsimony would argue for a single underlying cause that can be reflected in a single measure, the differences in thresholds and in the shapes of the psychometric functions in the present study provide a number of reasons to question such a conclusion. Furthermore, the present study has demonstrated that ABRM is not restricted to stimuli that are near the frequency DL.

ABRM is not a fixed attribute of auditory processing. Furthermore, the nature and variability in the shape of the psychometric functions generated in Experiments 2 and 3, as well as the psychophysical and psychometric functions of past ABRM studies, have exhibited "nonlinearities in $P(C)$, even in the middle range away from ceiling and floor effects" (Foyle & Watson, 1984, p. 521), and are not fit by simple quantitative models. These patterns of results are much more complex than would be expected from such a simple trading relationship or from a single source of interference.

In Experiment 2 (Figure 3, right panel), there were essentially two types of psychometric functions. The small

ABRM conditions exhibited a systematic increase in accuracy with increasing target tone duration up to approximately 36 msec, with very little change across longer target tone durations, and evidence for some persistence of interference even at the 96-msec target tone duration. The other two (H and L-H) maskers exhibited only a gradual change in accuracy over the range of target tone durations investigated and, even for the longest target duration, never reached the high level of accuracy of the small ABRM conditions. One of these maskers (L-H) was investigated in Experiment 3 as a function of masker delay (ISI), replicating the Experiment 2 (5-msec ISI) results. The 50-msec ISI psychometric function (L-H masker) became highly similar to the small ABRM functions (Experiment 2), where the ISI had been only 5 msec. The intermediate, 25-msec ISI psychometric function (Experiment 3) provides an indication that at least two processes may be involved, with accuracy similar to that of the 5-msec ISI function at the shortest target tone duration and accuracy levels similar to those of the 50-msec delay function (as well as those for the Experiment 2 small ABRM conditions) at longer target tone durations (e.g., above approximately 36 msec). Because the pattern of change at the low-accuracy end of the psychometric functions across the 5-, 25-, and 50-msec ISI masker delays involves a significant change in slope between 5- and 25-msec masker delays, it is not likely that the change in ISI is simply a leftward shift in function location. Furthermore, although its magnitude does change with increasing ISI, some interference is present at longer target tone durations. Thus, whether task difficulty is altered by varying target tone duration (present study) or ΔF (Foyle & Watson, 1984), as ISI is increased, the slope of the psychometric function changes and the upper tail remains depressed. The slope of the psychometric function reflects a trading relationship between task difficulty and ISI, and accuracy in its tail reflects the persistence of a relatively stable form of interference.

Although it is possible that the reduction in accuracy at long ISI values reflects disruption of pitch processes

on the occasional trials on which the needed processes are slow, it also is possible that it represents a different form of interference. Alternatively, this longer term disruption may involve interference in a later stage of processing, possibly at the decision stage, such as uncertainty about whether specific features arose from the target or the masker stimulus. It also might reflect some form of configural processing of features of the target and masking stimuli. A consideration that is relevant to both of these conjectures is that although the target stimulus in ABRM studies, in principle, is defined by temporal location, an accurate decision about the pitch of the target requires that an accurate temporal "label" be assigned to the sensory information that arose from the target stimulus. Such temporal labeling, however, probably does not occur until later stages in auditory perceptual processing (e.g., Cowan, 1988; Näätänen & Winkler, 1999).

Generality of Underlying Principles

The notion that perception is not instantaneous but, rather, progresses through a series of stages distributed across time is certainly not new, but the implications of the time lag, and the importance of stimulus properties and task demands to the time-delayed processes, often has not been adequately considered. The basic nature of these stages is consistent with the early, overlapping information processing stages posited by Näätänen and Winkler (1999) to account for the temporal pattern of evoked activity seen in the auditory cortex. Whether visual (e.g., Sperling, 1960) or auditory (e.g., Massaro, 1970) in nature, some form of sensory representation that decays in time is needed for poststimulus processing to occur. The initial processing probably identifies features, but the initial features probably lack an accurate temporal mapping (e.g., Näätänen & Winkler, 1999), with temporal attributes assigned as an aspect of a later stage that integrates features. One important addition to this common conceptualization is that both the importance and the efficiency of the poststimulus processing depends on the properties of the stimuli (e.g., duration, spectral separation) and the listener (i.e., skilled or unskilled). Although the temporal proximity of contextual stimuli may alter the sensory storage and processing stages, contextual stimuli also may influence the decision stage of processing. This information processing perspective, its components, and the auditory information principles that we have summarized are not unique to ABRM but should be basic properties of many types of auditory information processing. Since the present study showed evidence for ABRM for stimuli that are both longer in duration and more widely separated spectrally than those in past studies, the relevance of the underlying principles needs to be considered in interpreting perceptual results for more complex stimuli.

Temporal order judgment. The principles that result in ABRM should be relevant to defining the asynchrony duration threshold for identifying the order of onset of coterminous, spectrally distinct stimuli. The lower panel in Figure 1 illustrates a typical temporal order judgment (TOJ) paradigm, with the task being to identify which stimulus had the earlier onset. The TOJ task requires the listener to identify the brief segment of the stimulus that

is presented prior to the delayed onset of the other stimulus component. The fact that the spectrum of the earlier onset component continues in time and becomes more narrowband would be relevant only if the listener is able to both identify and temporally tag the features of the earlier component, the unique portion of the stimulus whose duration is the stimulus onset asynchrony (SOA). At the end of a brief SOA, the leading portion is broadband, making pitch identification difficult and forcing a reliance on later processing. In addition, what processing has been completed is probably still too early for assignment of a reliable temporal tag (e.g., Näätänen & Winkler, 1999). Finally, the portion of the stimulus after the SOA (i.e., when both stimuli are present) is not neutral, because, if that portion of the stimulus had no effect, the TOJ threshold would match the threshold for identifying a stimulus whose duration is equal to the SOA and, thus, approximately 8–10 msec (e.g., Doughty & Garner, 1947, 1948). However, Hirsh (1959) found that, for a wide variety of stimuli, the TOJ threshold is approximately 20 msec, and these findings motivated the studies that have demonstrated categorical perception along onset asynchrony continua, with 17- to 25-msec boundaries for spectrally distinct noise-buzz (Miller, Wier, Pastore, Kelly, & Dooling, 1976) and tone-pair (Pisoni, 1977) stimuli. It thus would appear that the latter portion of the stimulus in TOJ tasks interferes with the identification of the onset portion of the stimulus.

Further support for this conceptualization of TOJ and ABRM as reflecting the same underlying principles can be found in studies of the minimum component duration for identification of the order of consecutively presented tones. With relatively naive listeners judging the order of a fixed set of stimuli, threshold stimulus duration is 50 msec for three-tone sequences (Peters & Wood, 1973, cited in Divenyi & Hirsh, 1974) and longer for four-tone sequences (Warren, 1993). More important, the final stimulus in the sequence, being the only one not subject to ABRM, is consistently identified most accurately (Divenyi & Hirsh, 1974; Warren, 1993). Furthermore, in a three-tone order identification task, the addition of an invariant (fourth) tone that the listener is instructed to ignore results in a decrease in identification accuracy for the third tone, as well as for earlier tones in the sequence (Divenyi & Hirsh, 1975). Thus, the added fourth tone produces ABRM, essentially creating a perceptual version of the suffix effect (Crowder, 1986). Finally, although Hirsh (1959) found that the TOJ threshold is approximately 20 msec across a broad range of stimuli, longer TOJ thresholds have been reported for stimuli with dynamically changing onsets (e.g., Pastore & Farrington, 1996; Pastore, Layer, Morris, & Logan, 1988).

Psychophysical implications. Experiment 2 was motivated, in part, by the basic question of how to best characterize ABRM psychophysically. The present findings and conceptualization of ABRM and TOJ have implications not only for this basic question, but also for the common use of various psychophysical procedures. A number of psychophysical procedures (e.g., methods of limits and adjustment, the up-down adaptive procedure, the gating

paradigm) use systematic, sequential stimulus presentation to generate a single measure to reflect a processing limit (e.g., a threshold). One procedure used to study speech, the gating paradigm, evaluates the first point at which specific information first becomes available (e.g., Grosjean, 1996). In the typical version of this paradigm, as in the ascending method of limits, the duration of the stimulus is increased until accuracy in identifying the targeted feature, phoneme, or word reaches a criterion level. Implicit to this paradigm is the assumption that the subsequent portion of the stimulus, the portion not ever presented, is, at worst, neutral to the recognition task. The following hypothetical experiment, based on the present conceptualization of ABRM and TOJ, illustrates a potential problem with this and the other measures when this implicit assumption is invalid.

The top and middle panels in Figure 5 diagram a TOJ task—in this case, with the lower tone always leading. The broken vertical lines indicate the roughly 10-msec threshold for identifying the pitch of a brief segment of a sinusoid (e.g., Doughty & Garner, 1947, 1948) and the roughly 20-msec threshold for identifying the order of onset of a pair of tones (e.g., Hirsh, 1959). The two stimuli differ in SOA, which is 35 msec in the upper panel and 15 msec in the middle panel. Suppose that for each of these stimuli, the present procedure in Experiments 2 and 3 (method of constant stimuli) is used to map accuracy as a function of the duration of the stimulus—always beginning at stimulus onset, thus eliminating the ending portion of the stimulus. The expected (stylized) psychometric functions are illustrated in the lower panel in Figure 5. For the 35-msec SOA stimulus (solid line), as stimulus duration increases, accuracy should increase up to an asymptote (matching the no-masker psychometric functions in Experiment 2), then probably exhibit a small decrease after longer stimulus durations because the added lagging tone should introduce some residual (longer term) ABRM. For the 15-msec SOA stimulus, the psychometric function (broken line) should be identical up to 15-msec duration, after which all longer duration stimuli include the second tone. Since the 15-msec SOA for this stimulus is less than the approximately 20-msec TOJ threshold, the inclusion of the second tone should introduce ABRM interference that, with the SOA constant at 15 msec, should now be independent of stimulus duration. Thus, the psychometric functions for the two stimuli should be identical up to the 15-msec stimulus duration and should be very different at all longer stimulus durations.

If a typical gating paradigm (or the ascending method of limits) is used in this hypothetical experiment, the minimum stimulus duration for pitch identification should be approximately 10 msec for the stimuli at both SOAs (e.g., Doughty & Garner, 1947, 1948). The psychometric function for the 15-msec SOA stimulus would support a very different conclusion. Because processing of information is both distributed and delayed in time, the validity of the threshold estimated in this manner depends on whether the later portion of the stimulus significantly interferes with (i.e., 15-msec SOA), is neutral to (e.g., 35-msec SOA), or possibly even enhances the processing of earlier stimulus information.²

Figure 5. Illustration of a potential problem with gating and similar sequential paradigms. The upper two panels illustrate temporal order judgment (TOJ) stimuli with lower tone leading and respective stimulus onset asynchrony (SOA) values of 15 and 35 msec. The bottom panel plots expected psychometric functions. Broken vertical lines indicate the typical duration threshold for pitch identification and the typical SOA threshold for TOJ.

Possible generalization to natural stimuli. This example, and the present study, focused on simple stimuli (tones) that are widely separated in frequency. What are the implications for natural stimuli, such as speech, which are complex and often dynamic? If the perception of complex stimuli represents some form of Fourier synthesis, the processing demands for recognition are probably greater than those for the perception of the component tones. As a result, the minimum duration needed for recognition, and the dependence on poststimulus processing, is likely to be greater than that for tones. Some support for this conjecture is found in the amount of onset asynchrony needed to perceptually remove the formant from vowel and vowel-like stimuli, which ranges from 32 msec to longer than 90 msec, depending on the nature and duration of the harmonic stimulus (e.g., Hukin & Darwin, 1995). Furthermore, if, instead of being stationary, the stimulus component to be recognized is characterized by dynamic changes in frequency (e.g., formant transitions),

the greater complexity also increases the processing demands needed for recognition. Thus, although the present study (e.g., Experiment 1) indicated that the minimum duration for identification of an initial stimulus component is increased to 20–40 msec for tones far more distinct than the frequency DL, it is quite possible that for the recognition of early components of complex stimuli (such as in speech), the components need to be either longer (and possibly considerably longer) or more spectrally distinct. The possibility that spectrally dynamic onsets may be more susceptible to ABRM has been explored in a different study (Gaston & Pastore, 2008).

AUTHOR NOTE

This research was supported in part by National Science Foundation Grant BCS-0213666 to the first author. M.S.B. is now at the University of Toronto. The authors thank Jesse Flint for assistance in both the conduct of the research and the preparation of the manuscript. The assistance of many excellent undergraduate research assistants is also greatly appreciated. Correspondence concerning this article should be addressed to R. E. Pastore, Department of Psychology, Binghamton University (SUNY), Binghamton, NY 13902-6000 (e-mail: pastore@binghamton.edu).

REFERENCES

- EVERBACH, E., & CORIELL, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, **40**, 309-328.
- BLAND, D. E., & PERROTT, D. R. (1978). Backward masking: Detection versus recognition. *Journal of the Acoustical Society of America*, **63**, 1215-1217.
- COOPER, G. R., & MCGILLEM, C. D. (1967). *Methods of signal and system analysis*. New York: Holt, Rinehart & Winston.
- COWAN, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, **104**, 163-191.
- CROWDER, R. G. (1986). Auditory and temporal factors in the modality effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 268-278.
- DEARY, I. J. (1995). Auditory inspection time and intelligence: What is the direction of causation? *Developmental Psychology*, **31**, 237-250.
- DIVENYI, P. L., & HIRSH, I. J. (1974). Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America*, **56**, 144-151.
- DIVENYI, P. L., & HIRSH, I. J. (1975). The effect of blanking on the identification of temporal order in three-tone sequences. *Perception & Psychophysics*, **17**, 246-252.
- DOUGHTY, J. M., & GARNER, W. R. (1947). Pitch characteristics of short tones: I. Two kinds of pitch threshold. *Journal of Experimental Psychology*, **37**, 351-365.
- DOUGHTY, J. M., & GARNER, W. R. (1948). Pitch characteristics of short tones: II. Pitch as a function of tonal duration. *Journal of Experimental Psychology*, **38**, 478-494.
- DURLACH, N. I., MASON, C. R., GALLUN, F. J., SHINN-CUNNINGHAM, B., COLBURN, H. S., & KIDD, G. JR. (2005). Informational masking for simultaneous nonspeech stimuli: Psychometric functions for fixed and randomly mixed maskers. *Journal of the Acoustical Society of America*, **118**, 2482-2498.
- ELLIOTT, L. L. (1962). Backward masking: Monotic and dichotic conditions. *Journal of the Acoustical Society of America*, **34**, 1108-1115.
- FLETCHER, H. (1940). Auditory patterns. *Reviews of Modern Physics*, **12**, 47-65.
- FOYLE, D. C., & WATSON, C. S. (1984). Stimulus-based versus performance-based measurement of auditory backward recognition masking. *Perception & Psychophysics*, **36**, 515-522.
- GASTON, J. R., & PASTORE, R. E. (2008). *Non-simultaneous contrast effects in recognition for direction of frequency change with speech-like transitions*. Manuscript submitted for publication.
- GRIFFITHS, Y. M., HILL, N. I., BAILEY, P. J., & SNOWLING, M. J. (2003). Auditory temporal order discrimination and backward recognition masking in adults with dyslexia. *Journal of Speech, Language, & Hearing Research*, **46**, 1352-1366.
- GROSJEAN, F. (1996). Gating. *Language & Cognitive Processes*, **11**, 597-604.
- HAWKINS, H. L., & PRESSON, J. C. (1977). Masking and preperceptual selectivity in auditory recognition. In S. Dornic (Ed.), *Attention and performance VI* (pp. 195-212). Hillsdale, NJ: Erlbaum.
- HIRSH, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, **31**, 759-767.
- HOLDING, D. [H.], LOEB, M., & YODER, D. (1972). "Masking" versus interference in pitch perception. *Journal of Auditory Research*, **12**, 247-254.
- HUKIN, R. W., & DARWIN, C. J. (1995). Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification. *Perception & Psychophysics*, **57**, 191-196.
- KIDD, G. JR., MASON, C. R., DELIWALA, P. S., WOODS, W. S., & COLBURN, H. S. (1994). Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*, **95**, 3475-3480.
- 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.
- LUTFI, R. A., KISTLER, D. J., CALLAHAN, M. R., & WIGHTMAN, F. L. (2003). Psychometric functions for informational masking. *Journal of the Acoustical Society of America*, **114**, 3273-3283.
- MACMILLAN, N. A., & CREELMAN, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MASSARO, D. W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology*, **85**, 411-417.
- MASSARO, D. W. (1972). Stimulus information vs. processing time in auditory pattern recognition. *Perception & Psychophysics*, **12**, 50-56.
- MASSARO, D. W. (1977). Capacity limitations in auditory information processing. In S. Dornic (Ed.), *Attention and performance VI* (pp. 213-229). Hillsdale, NJ: Erlbaum.
- MASSARO, D. W., COHEN, M. M., & IDSON, W. L. (1976). Recognition masking of auditory lateralization and pitch judgments. *Journal of the Acoustical Society of America*, **59**, 434-441.
- MCARTHUR, G. [M.], & BISHOP, D. (2002). Event-related potentials reflect individual differences in age-invariant auditory skills. *NeuroReport*, **13**, 1079-1082.
- MCARTHUR, G. M., & HOGGEN, J. H. (2001). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *Journal of the Acoustical Society of America*, **109**, 1092-1100.
- MILLER, J. D., WIER, C. C., PASTORE, R. E., KELLY, W. J., & DOOLING, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. *Journal of the Acoustical Society of America*, **60**, 410-417.
- MOORE, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). New York: Academic Press.
- NÄÄTÄNEN, R., & WINKLER, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, **125**, 826-859.
- NEFF, D. L., & DETHLEFS, T. M. (1995). Individual differences in simultaneous masking with random-frequency, multicomponent maskers. *Journal of the Acoustical Society of America*, **98**, 125-134.
- OH, E. L., & LUTFI, R. A. (1998). Nonmonotonicity of informational masking. *Journal of the Acoustical Society of America*, **104**, 3489-3499.
- PASTORE, R. E., & FARRINGTON, S. M. (1996). Measuring the difference limen for identification of order of onset for complex auditory stimuli. *Perception & Psychophysics*, **58**, 510-526.
- PASTORE, R. E., LAYER, J. K., MORRIS, C. B., & LOGAN, R. J. (1988). Temporal order identification for tone/noise stimuli with onset transitions. *Perception & Psychophysics*, **44**, 257-271.
- PATTERSON, R. D., PETERS, R. W., & MILROY, R. (1983). Threshold duration for melodic pitch. In R. Klinke & R. Hartmann (Eds.), *Hearing, physiological bases and psychophysics. Proceedings of the 6th International Symposium on Hearing* (pp. 321-326). Berlin: Springer.

- PISONI, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops. *Journal of the Acoustical Society of America*, **61**, 1352-1361.
- ROBINSON, K., & PATTERSON, R. D. (1995). The duration required to identify the instrument, the octave, or the pitch chroma of a musical note. *Music Perception*, **13**, 1-15.
- RONKEN, D. A. (1971). Some effects of bandwidth-duration constraints on frequency discrimination. *Journal of the Acoustical Society of America*, **49**, 1232-1242.
- RONKEN, D. A. (1972). Changes in frequency discrimination caused by leading and trailing tones. *Journal of the Acoustical Society of America*, **51**, 1947-1950.
- ROSSING, T. D. (1990). *The science of sound* (2nd ed.). New York: Addison-Wesley.
- SPARKS, D. W. (1976). Temporal recognition masking—or interference? *Journal of the Acoustical Society of America*, **60**, 1347-1353.
- SPELTING, G. (1960). The information available in brief visual presentation. *Psychological Monographs: General & Applied*, **74**, 1-29.
- WARREN, R. M. (1993). Perception of acoustic sequences: Global integration versus temporal resolution. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 37-68). Oxford: Oxford University Press.
- 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.

NOTES

1. Threshold values also can be expressed in terms of number of cycles by dividing the results by 3.33, or in terms of the duration of one of the target tones, by multiplying the average threshold values by 0.8 for the high frequency or by 1.2 for the low frequency.

2. For the 15-msec SOA stimulus, whether or not there is convergence of thresholds for procedures that employ ascending and descending sequences (e.g., the up-down adaptive procedure) would depend on the starting durations for the different sequences.

(Manuscript received July 26, 2007;
revision accepted for publication March 17, 2008.)