1992

Lateralization of Asynchronous Stimuli

Mark A. Stellmack
Loyola University Chicago

LOYOLA UNIVERSITY CHICAGO

LATERALIZATION OF ASYNCHRONOUS STIMULI

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

MARK A. STELLMACK

CHICAGO, ILLINOIS

MAY 1992

ACKNOWLEDGMENTS

I would like to thank the director of my dissertation committee, William A. Yost, for his support and helpful feedback in the preparation of this dissertation and in the conduct of my research in general.

I would also like to thank Toby Dye, who guided me in my initial graduate studies. The early experiments that I conducted with him directly led to the present series of experiments. His feedback and support during the conduct of this research and preparation of this dissertation were of great value.

Thanks to Stan Sheft for his continuous input, guidance, and evaluation of my ideas.

I would also like to thank the other members of my dissertation committee for their comments and patience: Richard R. Fay and R. Scott Tindale.

VITA

The author, Mark Andrew Stellmack, was born March 24, 1964, in Chicago, Illinois.

In September, 1982, Mr. Stellmack entered Loyola University of Chicago, receiving the degree of Bachelor of Science in psychology in May, 1986.

In September, 1987, Mr. Stellmack was granted an assistantship in psychology at Loyola University of Chicago, receiving the degree of Master of Arts in psychology in May, 1990.

In September, 1991, Mr. Stellmack was granted a Schmitt Dissertation Fellowship at Loyola University of Chicago, enabling him to complete the Doctor of Philosophy in 1992. He is currently a member of the Acoustical Society of America.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Under normal circumstances, many different sound sources are simultaneously active in the environment of a human listener. The individual sounds produced by these sources interact and combine to produce a single complex sound pressure waveform at each of the listener's ears. Each sound source provides information that potentially allows the listener to identify <u>what</u> is producing a sound, <u>where</u> each sound-producing object is located, and the <u>meaning</u> or message in what is being conveyed by the sound-producing object. To extract this information from the complex sounds that arrive at each ear, the listener must be able to separate the complex sound pressure waveform into groups of spectral components that arose from common sources.

There are a number of cues that can give rise to the perceptual impression of one sound being segregated and distinct from other sounds that are present. (See Hartmann, 1988, and Yost, 1991a for reviews.) Spectral components that are harmonically related, have common modulations of amplitude or frequency, or have common onset times tend to be perceptually segregated from other components. Components that appear to originate from a unique spatial position relative to a background of additional sounds will also be perceived separately from the background. The lateral position of a sound source relative to a listener is based primarily on differences of time and level between the sounds arriving at each of the listener's ears. Sound sources in the external environment produce differences in time and level between the waveforms received by a listener's ears due to the geometry of the head. Interaural differences of time (IDT's) are generated when the sound source is located to one side of the listener so that the sound produced by that source reaches the ears at slightly different times. Interaural differences of level (IDL's) are produced when the sound arriving at the ear furthest from the sound source is attenuated by the head. These interaural cues are used to determine the location of the sound source in the azimuthal plane.

It is seemingly a common experience that an individual can identify the location of one sound source in the presence of many other sound sources. The direction from which a person's name is being called or from which a bus is bearing down on someone can often be identified in spite of other simultaneous environmental sounds. Such identifications require the listener to extract the correct interaural information associated with the sound of interest from conflicting interaural information. The present series of experiments examines the ability of listeners to detect the interaural differences of time (IDT's) of one or more spectral components in the presence of spectral components that bear conflicting interaural information.

When a pure tone is presented to a subject over loudspeakers, both IDT's and IDL's are present. As a result, it is difficult to separate the effects that these interaural parameters have on the listener's ability to locate the sound source, a task referred to as localization. IDI's and IDT's can be manipulated and studied independently by presenting stimuli over headphones. A pure tone presented over headphones that differs in intensity or time between the two ears (a "dichotic" pure-tone stimulus) is often described by the listener as producing a sound image "inside" the head, that is, at a particular intracranial position. When a listener attempts to locate the position of an intracranial image produced by a dichotic stimulus presented over headphones, the task is referred to as lateralization.

When two sound sources at different spatial positions emit spectrally non-overlapping signals, the stimuli received at a listener's ears consist of spectral components with different interaural delays and intensity differences. Stimuli in which different spectral components have different interaural information are described as spectrally incoherent. In order to accurately localize the individual sound sources, the binaural auditory system must be able to separate the conflicting interaural information and associate the appropriate interaural differences with their respective spectral components. However, Perrott (1984) showed that a listener is unable to identify the sources of two different pure tones emitted from different speakers as well as might be expected on the basis of the listener's ability to identify the sources of the tones individually. Perrott simultaneously presented two pure tones to subjects, each over a different speaker. He then measured the minimum angular separation required between the speakers in order for the subjects to identify which frequency component came from which speaker. He referred to this measure as the concurrent minimum audible angle. It was found that the concurrent minimum audible angles were several times larger than the minimum angle associated with identifying the source of the individual pure tones. Measurements taken by Kuhn (1977) show that between about 500 and 2000 Hz there is no unique relationship between interaural delay and azimuthal position, meaning that spectral incoherence can result from a single sound source containing energy between 500 and 2000 Hz.

Several experiments have shown that spectrally incoherent stimuli produce binaural interference across frequencies. Dye (1990) presented 3- and 5-component complexes to subjects over headphones and examined the effects of interaurally delaying different subsets of those components. In one experiment, stimuli were 3-component complexes centered at 750 Hz with component spacings ranging from 20 to 450 Hz. In the two-interval task, a subset of the 3-component complex was delayed to one ear in the first interval and to the other ear in the second interval, with the remaining component or components diotic across intervals. Subjects were instructed to make left-right judgments of the apparent movement of the delayed component(s), and threshold interaural differences of time (IDT's) were measured. Threshold IDT's were also measured for each of the three components in isolation. When only one component was delayed in the complex, thresholds were always elevated relative to that of the same component in isolation. In other words, the presence of the diotic components interfered with the ability of listeners to lateralize the delayed component. Thresholds were lower when two of the three components were delayed relative to the conditions in which only one of three components was delayed. Thresholds were approximately equal to those of a single component in isolation only when all three components in the complex were delayed. When the middle (750-Hz) component was delayed, thresholds were higher than when one of the outer components was delayed. In addition, there was an effect of component spacing, with the highest thresholds found at the 50-Hz spacing and decreasing thresholds at larger and smaller component spacings.

In a second experiment, Dye presented 3-component complexes in which two components were delayed to one side and the third component was delayed to the other side in the first interval of a 2IFC task, with the directions of the delays switched in the second interval. Subjects were instructed to report the apparent direction of movement of the complex across the two intervals. The proportion of responses consistent with the delay of the tone pair dropped continuously from 100% to 0% as the magnitude of the interaural delay of the third tone (that was delayed in the opposite direction) was increased. In other words, subjects' left-right judgments were jointly determined by the different interaural delays of the tone pair and the odd component.

These two experiments suggest that subjects fuse the components into a single intracranial image, with its apparent position a joint function of the relative interaural delays of the individual components. However, it has been shown for bands of noise that subjects report single intracranial images that broaden (instead of shifting position as a whole) with decreasing interaural coherence (Blauert and Lindemann, 1986), where interaural coherence is, in effect, the correlation between the waveforms presented to each ear. Only when the interaural coherence becomes sufficiently small do separate intracranial images result. Dye's (1990) results could be obtained if such broadening of the intracranial image occurred at small delays of a single component in a complex (relatively high interaural coherence). Larger delays (resulting in diminished interaural coherence) would be necessary in order for the images to split or become broad enough for a direction-of-movement judgment to be made. In Dye's second experiment, in which different components were delayed to different sides, subjects' judgments of the relative position of the intracranial image may have been based on the apparent lateral extent of a diffuse image.

A subsequent series of experiments (Stellmack, Dye, and Jakubczak, 1989; Stellmack, 1990) addressed the more basic issue of the sensitivity of the binaural auditory system to an interaural delay of a single component in a multi-component complex. In the majority of these experiments, listeners were presented a multi-component complex and asked whether or not the complex contained an interaural delay of the single target component. In contrast to Dye's experiments, listeners merely identified the presence of an interaural delay rather than the direction of the interaural delay on the basis of the lateral movement of the intracranial image.

In one of the experiments performed by Stellmack, Dye, and Jakubczak, the effects of the number of components and component spacing on the detection of an interaural delay of a 753-Hz target component were examined. Threshold IDT's were measured in the presence of 2, 4, 6, and 8 sinusoidal distractor components. The components were centered at 753 Hz with a frequency spacing of 10, 25, 50, 100, or 200 Hz. It was observed that threshold IDT's increased with increasing number of components. When other diotic components were present, threshold IDT's for the target component were always larger than those obtained for the target in isolation. For a given number of components, threshold IDT's were largest when the components were spaced 25-50 Hz apart and decreased at larger and smaller frequency spacings.

In another series of experiments, Stellmack and Dye (1989) manipulated a series of variables that were thought to aid segregation of the target from the distractor components. First, an onset asynchrony between the target and distractor components was introduced. The distractor components were gated on up to 200 ms prior to the target component to encourage perceptual segregation of the target component. Indeed, subjects reported that the pitch of the target component seemed to "stand out" against the diotic background, but threshold IDT's remained many times larger than those measured for the target component in isolation, and threshold IDT's were often many times larger than those measured in the first experiment, in which component onsets were synchronous. Woods and Colburn (1992) also observed that binaural interference occurred when an onset asynchrony between target and distractor components was introduced even though the pitch of the target became more salient.

Second, the harmonic relationship between the target and distractor components was varied, in addition to presenting an onset asynchrony between the components. It was observed that threshold IDT's for the target component were larger when additional components were present than when the target was presented in isolation regardless of the harmonic relationship of the components. In contrast, Buell and Hafter (1991) observed binaural interference between two pure tones only when the tones were harmonically related. No binaural interference was observed when the tones were inharmonically related.

In summary, Stellmack observed that threshold IDT's for a single target component in the presence of a number of additional diotic distractor components were larger than those for the target component presented in isolation. This result was obtained for up to 8 additional distractor components, frequency spacings from 10 to 200 Hz, and independent of the harmonic relationship between the target and distractor components.

In all of Stellmack's (1990) experimental conditions involving distractor components, the distractors were always present during the entire duration of the target component. Even in the conditions in which an onset asynchrony was present, the distractor components were gated on first and all components were gated off simultaneously. However, it is rarely the case in naturally-occurring situations that sounds from different sources will overlap so completely. Pure unmodulated sinusoidal stimuli are uncommon, as are stimuli from different sources with identical onsets and/or offsets. Rather, multiple sound sources produce sounds that differ in the rate and depth of both amplitude and frequency modulation in addition to differing in spectral content. The fact that sound sources in the real world most often produce sounds intermittently or with fluctuations in intensity means that, under normal circumstances, brief portions of the sound stimulus received by a listener consists of the sounds from the various sources in isolation. For example, in the simplest case of only two sound sources, A and B, if the stimulus produced by Source A contains a 50-ms temporal gap every 5 seconds (it is briefly and repetitively turned off), then the listener receives the stimulus from Source B in isolation for 50 ms at regular intervals. This isolated portion of Source B's stimulus might be sufficient to allow the listener to identify the location of the source.

The following series of experiments examines the effect of distractor components on the ability of a listener to detect the interaural delay of a target component when the components do not completely overlap over time. In the first experiment, the effects of temporal gaps of different durations in the diotic distractor components on the ability of a listener to detect the interaural delay of the target component will be examined. In the second experiment, two pure tones are rapidly alternated, or trilled, with one tone interaurally delayed in each of the listening intervals of each experimental trial. The ability of listeners to discriminate between delays of each tone will be measured. In the third experiment, the ability of listeners to detect an interaural delay of a target narrow-band noise in the presence of distractor noise bands will be measured when the target and distractors have either identical or different envelopes. As the duration of isolated presentation of the target becomes large, it is expected that performance will be similar to that for the target in complete isolation.

EXPERIMENT I

Threshold interaural differences of time (IDT's) of a 753-Hz target component were measured in the presence of six additional distractor components which were turned off briefly during the test interval of each experimental trial. The 7-component complex was centered at 753 Hz with a frequency spacing of 100 Hz. This number of components and frequency spacing were chosen because they resulted in substantial binaural interference in previous experiments (Stellmack, 1990). Thresholds were measured as a function of the temporal notch in the distractor components. Subjects judged whether or not the 753-Hz component of the complex was interaurally delayed on each trial. As the duration of the temporal notch increases, thresholds should approach those measured for the target in isolation. This will give an indication of the duration of isolated presentation required to eliminate the interference observed in previous experiments.

METHODS I

A cued single-interval task was used in which a diotic 753-Hz pure tone was presented (the "cue tone") followed by the test stimulus. Thus, each trial consisted of two presentations in which only the second varied from trial to trial. The cue tone was intended to help subjects identify the intracranial center and to help them attend to that particular pitch in the test interval. The target component was always 753 Hz and the distractors, present only during the test interval, were 453, 553, 653, 853, 953, and 1053 Hz. The duration of the intervals depended on the particular condition being run as described below. The two intervals of each trial were separated by 300 ms of silence and all components were given 10-ms linear rise/decay times.

The test interval consisted of either a diotic complex or a complex in which only the 753-Hz component was delayed to the right ear. This task is referred to as a "left-center" task because the interaural delay, when present, results in an image to the left of the midline for a pure tone, while a diotic pure tone appears to be centered. All components were equal in amplitude (55 dB SPL) and all components were gated simultaneously, thus interaural delays of the 753-Hz component were achieved by advancing the phase of that component in the left channel. The starting phases of all components in the test complexes were randomized between trials to eliminate monaural cues that can result from shifting the phase of one component in the complex relative to the others. The observer's task was to indicate, by pressing one of two response buttons on a terminal, whether or not the 753-Hz component was delayed during the test interval. Feedback was given to the subjects on a trial-by-trial basis.

In one set of conditions, a temporal notch in the distractors was introduced 200 ms into the test interval. Threshold IDT's of the target component were measured for notch durations of 0, 10, 25, 50, 100, and 200 ms. (In the 0-ms notch duration condition, the distractors remained on for the entire test interval.) The target component was on for the entire 500-ms test interval. Linear decay and rise times of 10 ms were used to produce the temporal notch in the distractors. The duration of the notch refers to the time during which the distractors were completely turned off, not including the duration of the linear gating of the distractors. This set of stimulus conditions is depicted in the top portion of Figure 1.

A second set of conditions slightly different from that described above was also run, with the main difference being that the target component was present only during the temporal notch in the distractors. (See the bottom portion of Figure 1.) Once again, the temporal notch in the distractors was introduced 200 ms into the test interval. When the distractors were completely off, the target component was turned on for the duration of the notch, and then the distractors were turned on again for the remainder of the 500-ms test interval. In these conditions, the target and distractors were completely non-overlapping. Linear rise/decay times of 10 ms were used for the gating of all components. Threshold IDT's were obtained for notch durations of 50, 100, and 200 ms in this set of conditions. Shorter durations were not run because the duration of the onset and offset gating functions would exceed that of the stimulus itself. Shorter gating functions would produce excessive smearing of the stimulus spectrum.

Thresholds were also obtained for a 753-Hz tone in isolation, in which case only that tone was present in the test interval. Threshold IDT's were measured for target durations of 50, 100, 200, and 500 ms.

In all of the above conditions, the cue tone (which consisted of the target component alone) was of the same duration as the target component in the test interval. In the first set of conditions, in which the target was on for the entire test interval, the cue tone was 500 ms in duration. In the second set of conditions, in which the target was on only during the temporal notch in the distractors, the duration of the cue tone was equal to the duration of the notch, which also equalled the duration of the target in the test interval. The same is true for the conditions in which thresholds were measured for the target in isolation.

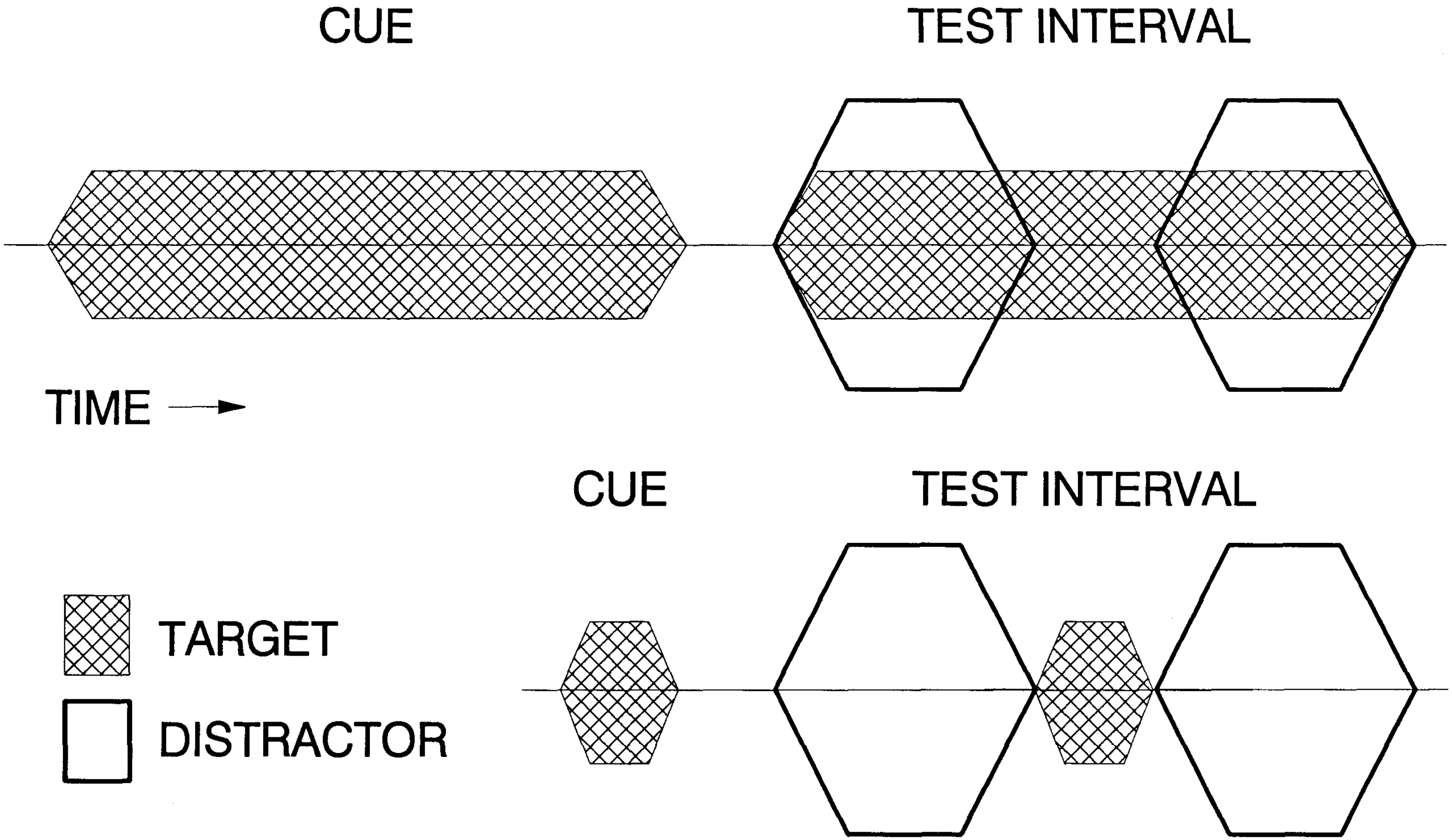
Thresholds were estimated from 3- or 4-point psychometric functions, with each point based upon at least 100 trials. It was desirable for the subject's first block of trials to have a relatively large delay. When a subject was run on blocks of trials for a particular condition on more than one day, it was occasionally necessary to run a block of trials at an IDT that was run the previous day. In these cases, data was combined for the blocks of trials with the same IDT. The threshold interaural delay was defined as the delay estimating a d' of 1.00 by linear interpolation.

Stimuli were presented through Telephonics (TDH-49) earphones suspended in Auraldomes to subjects seated in a sound-attenuating chamber. All stimuli were generated and presented by a Masscomp minicomputer interfaced with 16-bit digital-to-analog converters whose output rates were set to 20 kHz per channel. The signals were low-pass filtered at 7.5 kHz for antialiasing (Rockland Series 2000). The levels of the signals were adjusted with variable attenuators (Tech Lab, Inc.) before being passed on to Crown stereo amplifiers (Power Line Two) which were used to drive the earphones.

Data were gathered in blocks of 100 trials, with each set of 50 trials separated by a short rest period. Before each set of 50 trials, subjects were allowed to listen to practice trials which were like those to be presented during the experimental run. Subjects were instructed to adjust the position of the headphones during practice trials so that the intracranial image produced by the diotic cue tone appeared to be centered. When ready, subjects initiated a set of experimental trials by pressing a button on the response terminal. Data were gathered during two-hour sessions during which subjects were run individually. A typical experimental session consisted of 500-600 trials per subject.

The three observers in Experiment I were the author and two undergraduate volunteers from the author's university who were paid an hourly wage for their participation. All of the listeners had extensive experience from participation in other lateralization studies. As a result, a minimal training period was required to familiarize the observers with the task and the type of stimuli to be lateralized.

Figure 1. Depiction of the target and distractor envelopes in two of the stimulus conditions presented in Experiment I. The cross-hatched portions represent the target and the bold outlines represent the distractors. The top portion of the figure represents the stimulus condition in which the target component was on for the entire test interval. Both the cue and test interval are 500 ms in duration. The bottom portion of the figure represents conditions in which the target component was on only during the temporal notch in the distractors. The cue duration is equal to the duration of the target component of the observation interval while the entire test interval is 500 ms in duration. The linear gating of components is not drawn to scale.


CUE
TEST INTERVAL
TIME
CUE
TEST INTERVAL
TARGET
DISTRACTOR

RESULTS I

Figure 2 shows the results for each of the three subjects. For the open symbols, threshold IDT's are plotted against the duration of the temporal notch in the distractors. The circles represent data from the first set of conditions, in which the target was on for the entire duration of the observation interval (500 ms). The squares represent data from the second set of conditions, in which the target was on only during the temporal notch in the distractors. Threshold IDT's are plotted against stimulus duration for the stars, which represent conditions in which the target was presented in isolation. The dotted horizontal line near the bottom of each panel is the subject's threshold for a 500-ms 753-Hz tone in isolation. Each panel shows data from a single subject. Note that the ordinate of the graph for Subject 2 is scaled differently than those for the other two subjects in order to accommodate all of the data points.

Consistent with Stellmack's (1990) observations, it can be seen that the presence of diotic components interferes with the ability of listeners to detect the interaural delay of the 753-Hz component. All thresholds for the complex stimuli are higher than those for the 753-Hz tone in isolation, although they are only slightly so for the longest notch duration.

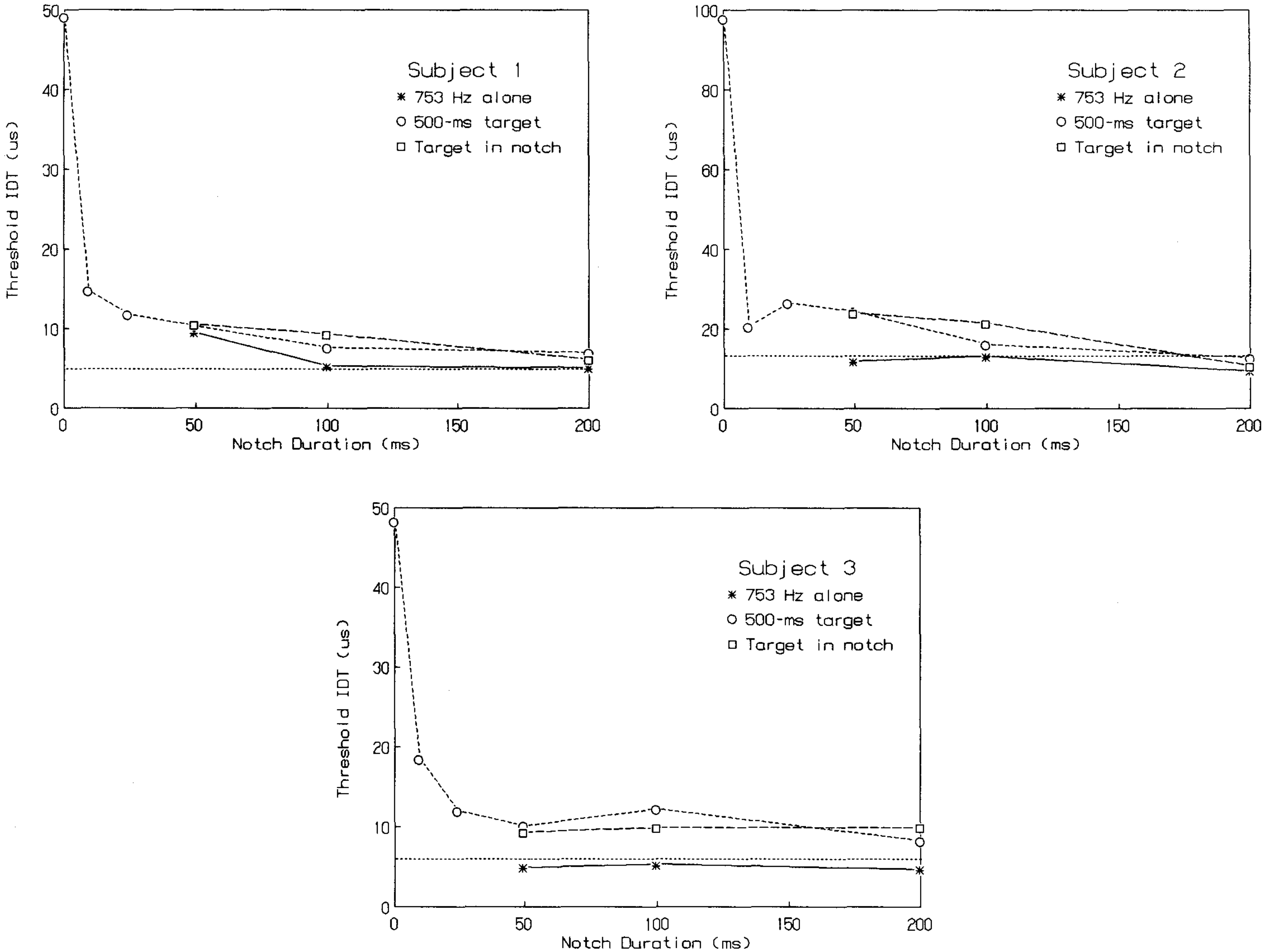
A brief temporal notch in the distractors had a large impact on the ability of listeners to detect the interaural delay of the target component. Thresholds measured for the completely overlapping target and distractors (open symbol, 0 ms notch duration) are approximately 4 to 5 times as large as thresholds measured with distractor notches of 25-200 ms. Subjects 1 and 3 show a decreasing ability to detect the IDT of the target component for a distractor notch of 10 ms: thresholds for the completely overlapping target/distractors are only about 2 1/2 to 3 times as large as those measured in the 10-ms notch condition. The threshold for Subject 2 in the 10-ms notch condition is actually lower than that for some of the longer durations. Presumably, this subject's thresholds would begin to rise in a manner similar to that of the other subjects at shorter durations.

In comparing thresholds measured when the target is on only during the distractor notch (squares) to thresholds measured when the target is on for the entire test interval (circles), it can be seen that the 500-ms target gives the listener no advantage in detecting the target IDT. Thresholds in these two sets of conditions are nearly equal.

When a distractor notch is present, thresholds approach those measured for the target in isolation. Except for the 50-ms notch threshold for Subject 2, thresholds for the distractor notch conditions (open symbols) differ from the thresholds measured for the target in isolation (stars) by less than 10 μs for notch durations of 50-200 ms.

A minimal effect of target duration on thresholds was observed for the target presented in isolation across the durations that were examined (50-500 ms). When the target was presented in isolation, thresholds were within 5 μs of one another across all durations for each subject.

Figure 2. Threshold interaural differences of time (in μs) are plotted as a function of the distractor notch duration (in ms) for the open symbols. Circles represent conditions in which the target was on for the entire observation interval, squares represent conditions in which the target was on only during the notch. For the stars, threshold IDT's are plotted as a function of stimulus duration (in ms), as these symbols represent thresholds for the target in isolation. The dotted horizontal line is the subject's threshold for a 500-ms target in isolation. The target was 753 Hz and the distractors, when present, were 453, 553, 653, 853, 953, and 1053 Hz. Each panel in the figure shows data from a single subject.


Subject 1
753 Hz alone
500-ms target
Target in notch
Threshold IDT (us)
Notch Duration (ms)
Subject 2
753 Hz alone
500-ms target
Target in notch
Threshold IDT (us)
Notch Duration (ms)
Subject 3
753 Hz alone
500-ms target
Target in notch
Threshold IDT (us)
Notch Duration (ms)

DISCUSSION I

When a notch in the distractors is present, the results suggest that only the information that appears during the distractor notch is useful in lateralizing the target. Thresholds for the 500-ms target conditions are quite similar to those for the target-in-notch conditions. There is no advantage gained in lateralizing the target by having the target on for the entire test interval. This is to be expected because the threshold IDT's for the notched-distractor conditions are well below those for the no-notch condition, so presumably the IDT of the target is completely undetectable while the distractors are present.

The smallest temporal notch used in this experiment, 10 ms, was sufficient to lower thresholds substantially relative to the no-notch condition, and notches of 25 ms or more produced thresholds within 10 μs of those measured for the target component in isolation for Subjects 1 and 3, and within 13 μs for Subject 2. These results suggest that the binaural interference produced by the distractor components decays to a large extent within 10 ms and almost entirely by approximately 25 ms after distractor offset.

Perhaps in some ways the present experiment is similar to those examining the effects of changes in the ongoing configuration of a stimulus on the masking level difference (MLD). In most MLD experiments, detection thresholds are measured for a tone in noise under circumstances in which the tone and noise have different interaural configurations. The tone, or signal, is identified by S with a subscript denoting the interaural phase configuration of the tone. For example, S_π indicates a signal with an interaural phase shift of 180°. Similarly the noise masker is identified by N with a subscript indicating its interaural phase configuration. Under circumstances in which thresholds for an S_π probe tone are measured in the presence of an N_0 masker, thresholds are approximately 15 dB lower than when thresholds are measured for the same signal in the presence of an N_π masker. The difference between the N_0S_π threshold and $N_\pi S_\pi$ threshold is the masking level difference (MLD).

McFadden (1966) and Yost (1985) showed that the magnitude of the MLD is as much as 13 dB larger when the noise masker is continuous than when the noise and signal are gated on simultaneously. In addition, Yost (1985) showed that MLD's also increase when the noise is gated on from 100 to 500 ms prior to the signal, with MLD's increasing as the difference in time between the noise and signal onsets increases. Because the signal and noise have different interaural configurations in the N_0S_π condition, presumably the noise alone and the signal-plus-noise stimuli occupy different intracranial positions. When the signal and noise are gated on simultaneously, the listener must base his judgments on the absolute position of the stimulus. On the other hand, if the noise is gated on first followed by the signal (in signal trials), <u>movement</u> of the intracranial image occurs during the stimulus interval. These results suggest that a moving stimulus is easier to detect than a stationary one, as long as the listener has adequate time to determine the position of the noise background (at least 500 ms for the maximum effect) before the signal is introduced and movement of the image occurs.

Yost, Turner, and Bergert (1974) presented evidence suggesting that detection of movement is a fundamentally different, and easier, task than that of detecting the absolute position of a stimulus. In one condition, listeners attempted to detect the interaural delay of a pure tone in a yes-no task, in which a single tone was presented to subjects either diotically or with an interaural delay. A same-different task was also presented in which two intervals were presented to listeners, the first interval always containing a diotic pure tone and the second interval containing a diotic tone or one with an interaural delay. If both tasks are performed by determining the position of the stimulus, they should produce equal detectability for a given IDT according to the theory of signal detection, because both tasks contain only one interval with useful information. Yost et al. (1974) found that the two methods did not produce similar results. Performance on the same-different task was superior to that on the yes-no task because listeners had a movement cue in the same-different task that was not present in the yes-no task.

The results of the present experiment might be explained in terms of the introduction of a movement cue in those conditions where the distractors are turned off briefly. When the distractors remain on for the entire listening interval, no movement of the image occurs because the interaural configuration of the stimulus remains constant throughout the entire listening interval. However, the present experiment differs from the MLD experiments described above in that the MLD experiments involve an introduction of both the target pitch and its apparent position after some duration of the noise alone, and the listener's task is to merely detect the presence of the target. In the present experiment, the target is always present and only the interaural delay is introduced when the distractors are turned off. In previous work by the author that is more comparable to the MLD experiments described above (Stellmack, 1990), an onset asynchrony between the target and distractors did not make interaural delays of the target easier to detect, although they did make the pitch of the target stand out. In addition, the interaural delays of the target used in the MLD experiments described above are much larger than those used in Experiment I. This suggests that the effects of a delayed onset of the target and an extremely large interaural delay combine to make the target more detectable, but does not address the issue of detectability of the interaural delay itself.

Previous research has demonstrated the "sluggishness" of the response of the binaural system to changes in the interaural configuration of a noise stimulus (Grantham and Wightman, 1978; Kollmeier and Gilkey, 1990). Grantham and Wightman (1978) examined the detectability of a varying interaural difference of time of a broadband noise stimuli. They generated two frequency-modulated noises with opposite phases of modulation. Because the noises were modulated out of phase, the fine structure of one noise alternatingly led and lagged that of the other during one period of modulation. When two noises generated in this way are presented to opposite ears, they produce an intracranial image that moves sinusoidally back and forth between the ears. Grantham and Wightman asked subjects to discriminate this moving stimulus from a stationary one, in which the same FM noise was presented to both ears. Thresholds were measured in terms of the peak IDT needed in the moving stimulus to discriminate it from the stationary one. Thresholds were measured for a range of modulation frequencies (rates of movement) from 2.5 to 500 Hz.

Grantham and Wightman found that threshold peak-IDT's increased steadily as the modulation frequency (f_m) increased from 2.5 to about 50 Hz. Threshold peak-IDT's decreased steadily as f_m increased above 50 Hz. Listeners reported that they detected the movement of the stimulus for $f_m < 50$ Hz, and they based their decisions on the apparent breadth or diffuseness of the intracranial image for $f_m > 50$ Hz. Thus, the ability of listeners to detect the movement of the stimulus decreased as the movement rate increased. The authors interpret their data as indicating that the binaural system displays a low-pass characteristic with a cutoff frequency of about 10 Hz with respect to fluctuating binaural input. The binaural information of a stimulus that fluctuates at a faster rate is "smoothed" in terms of its apparent instantaneous position as a result of the binaural system's inability to follow the rapid interaural fluctuations of the stimulus, with the effect of reducing the perception of movement produced by these fluctuations.

In order to compare the results obtained with the sinusoidally fluctuating IDT of Grantham and Wightman to those obtained with the essentially rapidly switched IDT of the present experiment, which is more similar to a step function, some way to equate the two stimuli must be found. Because the stimulus of Grantham and Wightman contains sinusoidally-fluctuating IDT's, and keeping in mind that the stimulus alternatingly leads to each ear, the stimulus will lead to each ear by greater than or equal to one-half the peak IDT for one-third of the period of modulation. For example, if the instantaneous IDT begins at 0 and changes such that the stimulus first leads to the left ear, the instantaneous IDT will exceed one-half the peak IDT between 30° and 150° in the phase of the modulator and the resulting intracranial image will be to the left of the midline. The instantaneous IDT will again exceed one-half the peak IDT between 210° and 330°, but with the intracranial image to the right of the midline. For a modulation frequency of 10 Hz, having a period of 100 ms, these fractions of a period correspond to approximately 33 ms during which the IDT exceeds one-half the peak IDT in the direction of each ear.

If we arbitrarily assume that the effective integration time of binaural information is the duration during which the instantaneous IDT exceeds one-half of the peak IDT, performance in the Grantham and Wightman study begins to drop off at effective burst durations of dichotic information that are 33 ms (for f_m = 10 Hz). In fact, performance in Grantham and Wightman's study falls off steadily with increasing f_m from f_m = 2.5 Hz. This is inconsistent with the present experiment in which performance is essentially flat for notch durations greater than 20 ms, and performance drops off only slightly between 10 and 20 ms. Considering that listeners in Grantham and Wightman's study received multiple looks at the interaurally delayed stimulus (twice each period of modulation), one might expect even better performance than in the present experiment, but this is not the case. Certainly, the important difference between these experiments must be that the IDT was smoothly changed by Grantham and Wightman, while in the present experiment, the IDT of the target was abruptly revealed during the distractor notch.

Perhaps the present experiment is more comparable to a study of MLD's by Kollmeier and Gilkey (1990). In their study, detection thresholds were measured for a 20-ms, 500-Hz, interaurally phase-inverted probe tone (S_π) in the presence of a noise masker which was abruptly altered in one of two ways during its presentation. In one set of conditions, the noise masker was abruptly changed from interaurally phase-inverted (N_π) to interaurally in-phase (N_0). In a second set of conditions, the interaural phase of the noise masker was held constant (N_π) but the level of the masker was reduced by 15 dB. The detection threshold for the probe tone was measured as a function of its temporal position relative to the point in time at which the noise masker was changed.

In their experiment, Kollmeier and Gilkey observed that when the S_π signal was presented after the noise masker was switched from N_π to N_0, thresholds gradually and continuously decreased to the expected N_0S_π level as the time between the switch and signal presentation increased to 200 ms. In the second set of conditions, thresholds decreased continually as the time between a 15 dB masker level decrease and signal presentation increased. The decrease in thresholds occurred much more rapidly (over 100 ms) when the level of the masker was changed than when the interaural configuration of the masker was changed. When the interaural phase of the masker is changed, the effective level of the masker is not altered in monaural channels, so changes in performance are due to the binaural system. The results were interpreted as an indication that the binaural system reacts more "sluggishly" to temporally varying stimuli than the monaural system.

Although Kollmeier and Gilkey utilized an abruptly changing stimulus configuration as in the present experiment, significant differences between the two experiments still exist. Listeners in Kollmeier and Gilkey's study attempted to detect the presence of a signal that had a fixed interaural phase difference, while in the present experiment, listeners tried to detect the interaural delay of a target with a fixed intensity level. Kollmeier and Gilkey observed a relatively slow response of the binaural system to changing interaural information similar to that seen by Grantham and Wightman. The present experiment demonstrates a situation in which the binaural system responds to very brief changes in the interaural configuration of the stimulus.

Suppose that when a new auditory stimulus is introduced, a description of that stimulus is formed and entered in what might be called an "auditory descriptor buffer" in memory. This buffer might contain a description of each currently active stimulus detected by the system. Each description would include all relevant information about that stimulus, such as the frequencies that are most likely part of that stimulus, information as to modulation or temporal patterns of those frequencies, as well as the position of the sound source producing the stimulus. (Certainly other types of information might be included.)

Furthermore, imagine that this description is very quickly formed when the stimulus is first introduced and that the information in that description is updated as the stimulus continues. It may be that it is difficult to modify information in this description once it is formed (perhaps it is weighted more heavily than new conflicting information), so it would take time for incoming information as to the sound's current position to outweigh and supercede that existing in the description. On the other hand, it is possible that the information in the stimulus description is updated on a priority basis, based on the relative importance of different elements of the description. In terms of comprehending speech, for example, it is probably more important to carefully and accurately follow the frequency and timing fluctuations of the stimulus than its apparent position, so the direction of the sound source might be updated less frequently. In addition, if a sound source is of interest to a listener, localization of the sound source is usually followed by orientation toward the sound source and perhaps visual contact, making subsequent localization superfluous.

It is possible that the distinguishing factor between the present experiment and those of Grantham and Wightman (1978), and Kollmeier and Gilkey (1990) is that these other experiments involve changes in the interaural configuration of what might be considered to be an existing auditory object. In the case of Grantham and Wightman, the interaural delay of an ongoing noise stimulus is gradually changed. In the Kollmeier and Gilkey study, the noise is turned on with a particular interaural delay which changes at some point during the stimulus presentation. If the concept of a description of the stimulus formed in memory as described above is accurate, updates of the binaural information associated with the ongoing stimuli are difficult or have low priority. Perhaps in the present experiment, when the target and distractors are turned on together, the target component is perceptually fused with the distractor components, forming a single auditory object and a single description of that stimulus in memory, complete with its apparent position. When the distractors are turned off, the target becomes established as a new and unique auditory object in the perceptual field, separate from the distractors, which prompts the formation of a new description of the stimulus. If formation of a new stimulus description is assumed to occur quickly, the binaural information associated with the target is entered in memory very quickly after the distractors are turned off.

The type of model described above would also account for the results of the MLD experiments described earlier (McFadden, 1966; Yost, Turner, and Bergert, 1974; Yost, 1985). In those experiments, when the signal is turned on after some duration of noise presentation, the introduction of a new auditory stimulus prompts the construction of a description of that stimulus. There are two separate cues to the auditory system that a new event has occurred: the signal has a pitch quality different from that of the noise background and it appears to occupy an extremely lateral position relative to prior stimulation. When the signal has no IDT, only the pitch cue is present to indicate the occurrence of a new event.

The model described above does not preclude the notion that interaural information is processed across frequencies in a spectrally-synthetic manner. When a new auditory event occurs in the presence of ongoing stimuli, a new entry may be formed in memory for that stimulus, but interference across frequencies might prevent the interaural information associated with the new event from being perceived accurately. The result would be the perception of a unique auditory event in terms of its pitch, for example, but at an inaccurate spatial location. This describes the situation that occurred in Stellmack (1990) when the target component was introduced after an onset asynchrony in the presence of several distractor components. The target was readily perceived as a unique event in terms of its pitch, but it was difficult for subjects to detect the interaural delay of the target.

EXPERIMENT II

When two pure tones are rapidly alternated, or trilled, a listener will most likely perceive the stimulus in one of two ways. The listener will either report a single stream of auditory stimulation consisting of two alternating tones, or the listener will report hearing two different sequences of repeating tones occurring at the same time. (See Bregman, 1990, and Handel, 1989 for reviews of the literature on stream segregation of rapidly alternated tones.) In the latter case, even though the tones do not actually temporally overlap in terms of the acoustic waveform received at a listener's ear, the listener perceives two different but simultaneous events, namely, two repeating series of pulses with different frequencies. The listener also reports that he or she is able to switch attention from one stream of tones to the other. Whether the stimulus will be perceived as one stream or two is partially determined by the frequency spacing between the tones and the rate at which the tones are alternated. In general as the trill rate increases and/or the frequency spacing increases, it becomes more difficult for the listener to perceive the tones as a single stream. (The tendency for rapidly alternated tones to be perceived as separate streams was understood and utilized by composers centuries ago to achieve the impression of two melodies being simultaneously played by a single instrument.) In many cases, when the trill rate and frequency spacing are at moderate levels, the listener is able to choose between the perception of a single stream and two concurrent streams, in much the same way an observer can alternatingly perceive portions of a visual stimulus as figure or ground in certain ambiguous visual displays. In this way, attention plays an important part in determining the manner in which the stimulus is perceived.

If the alternating tones are also played to different ears over headphones, the additional cue of the spatial separation of the tones further encourages the listener to perceive the repeating tones as two separate events. Judd (1979) observed that when different patterns of tones were presented to each ear with the individual tones alternating between the ears, listeners could correctly identify the order of the tones in each ear, but could not identify the order of the tones between the two ears. For example, suppose the following stimuli were presented simultaneously to each ear (The numbers 1 through 4 indicate different pitches and * indicates silence.):

Left: 1 * 4 * 1 * 4 ...etc.

Right: * 2 * 3 * 2 * ...etc.

Thus, when a tone is played to the left ear, there is silence in the right headphone channel and vice versa; that is, the tones are presented in isolation. A tone presented to only one ear will usually be lateralized at an intracranial position at the ear of presentation. In the example described above, listeners can easily identify which tones are presented to each ear, but they cannot accurately name the order of tones across channels, for example, whether pitch 3 follows pitch 1 or 4. This result indicates that listeners have access to the apparent positions of the individual tone pulses, because it was on the basis of their apparent position that the stimulus was organized into streams. Two unique, independent auditory events are heard. When Judd replaced the silent portions of the stimulus described above with broadband noise, the lateralization of the tones toward the ear of presentation was reduced and the series of tones was perceived as a single perceptual stream, with the result that listeners were able to correctly identify the order of the presentation of tones across ears.

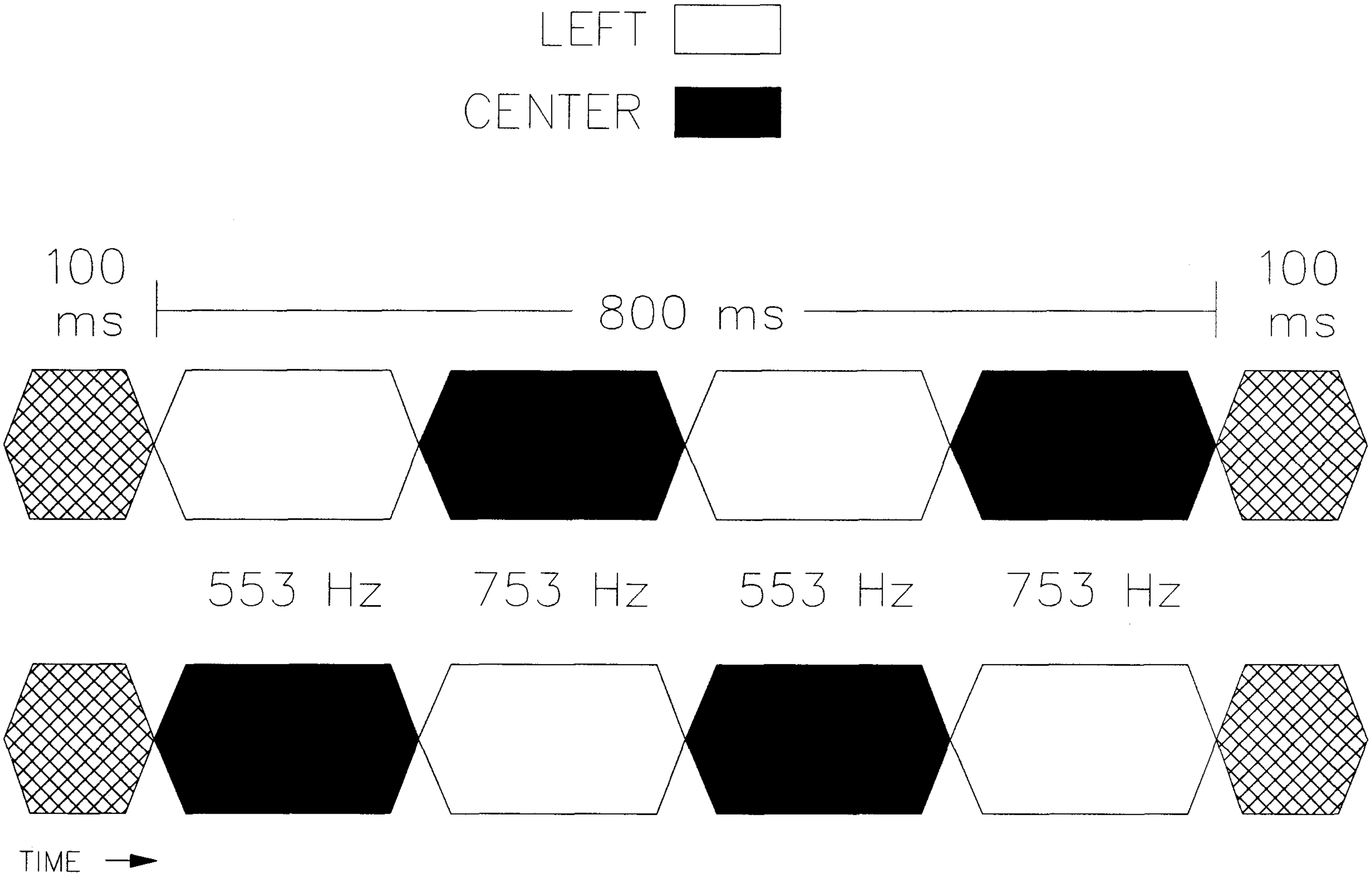
Experiment II will examine the ability of listeners to discriminate between stimuli consisting of simultaneous auditory events with different binaural information. The simultaneous auditory events in this case are two rapidly alternating sinusoidal components of different frequencies in which the individual tone pulses do not temporally overlap. Unlike Judd (1979), spatial separation of the tones will not be produced by presenting the tones to only one ear, but by introducing an interaural delay to the pulses of one frequency and not to the other. The tones will be sufficiently distant in frequency and presented rapidly enough to permit segregation on the basis of pitch. It is expected that the presence of cues promoting the perception of separate auditory streams based on pitch (spectral separation of components, temporal asynchronies between pulses), combined with the fact that the components are briefly presented in isolation during the course of the stimulus, will allow listeners to lateralize the pitch streams accurately, with little interference across frequencies.

METHODS II

Figure 3 depicts the two intervals of each trial in this experiment. Each interval was 1000 ms in duration, separated by 300 ms of silence. The first and last 100 ms of each interval consisted of a 553-Hz tone and a 753-Hz tone presented simultaneously and diotically. During the middle 800 ms of each interval, the 553-Hz tone and 753-Hz tone were rapidly alternated, with the 553-Hz tone pulsed on first. The 553-Hz pulses were interaurally delayed to the right in one interval with the 753-Hz pulses presented diotically. In the other interval, the 553-Hz pulses were diotic and the 753-Hz pulses were interaurally delayed to the right. As a result, if each stream of pulses could be heard at a separate intracranial position, a stream of pulses of one frequency should appear to be in the center of the listener's head and the stream of pulses of the other frequency should appear to be to the left during each interval. The interaural difference of time was equal across intervals within each trial. The 100-ms diotic pulses were placed at the beginning and end of each interval to eliminate any advantage there might be in lateralizing the first 553-Hz pulse or last 753-Hz pulse of the observation interval. All pulses were gated on and off with a 5-ms raised cosine function.

In one set of conditions, the pulsed tones were completely non-overlapping in time (shown in Figure 3). Threshold IDT's were measured for pulse durations of 20, 50, 100, and 200 ms. The number of pulses

Figure 3. A depiction of the non-overlapping stimulus presented in Experiment II. The top and bottom portions of the figure represent the first and second intervals of an experimental trial. The first and last 100-ms pulses (cross-hatched) were diotic complexes consisting of both 553-Hz and 753-Hz sinusoidal tones. During the middle 800 ms of each interval, 553-Hz and 753-Hz pulses were alternated, with their apparent positions (indicated by light and dark shading) changing between intervals. The figure represents one trial of the 200-ms non-overlapping pulse condition. Raised-cosine gating functions were used to gate all pulses on and off in the experiment, but are not represented in the figure.


LEFT
CENTER
100 ms
800 ms
100 ms
553 Hz
753 Hz
553 Hz
753 Hz
TIME

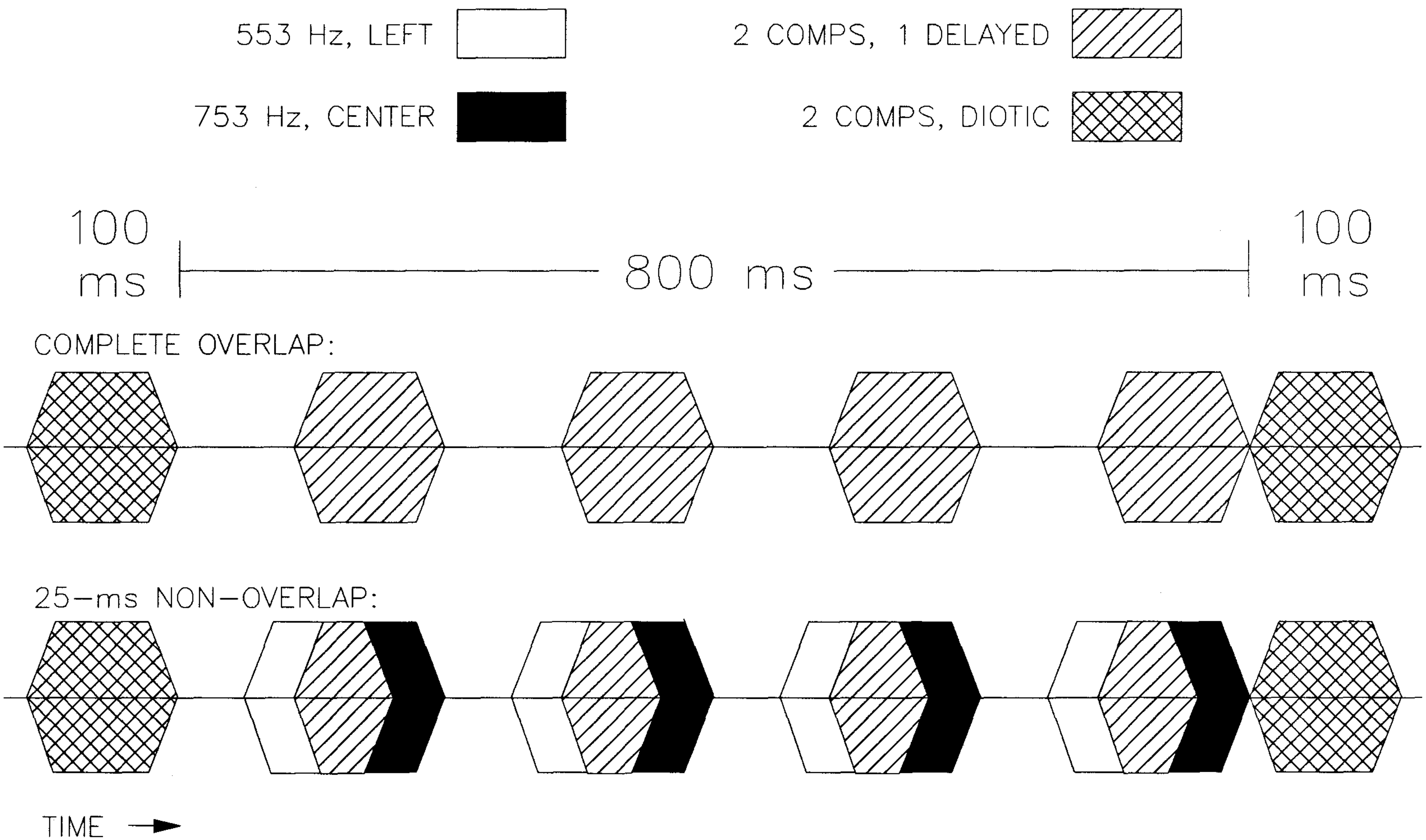
during each interval depended on the pulse duration. For example, during the middle 800 ms of each interval at the 20-ms pulse duration, 40 pulses were presented, 20 of each frequency. At the 200-ms pulse duration, only four pulses were presented, two of each frequency. (In all of the conditions of this experiment, pulse duration is perfectly confounded with number of pulses, but the total presentation time of each frequency is constant at 400 ms.)

In a second set of conditions, the stream of 553-Hz pulses was shifted in time by the duration of one pulse so that it completely overlapped the stream of 753-Hz pulses during the middle 800 ms of each interval. (See the top portion of Figure 4.) Thus, on a trial of the 50-ms pulse condition, the initial 100-ms diotic burst was followed by 50 ms of silence, then a 50-ms pulse consisting of both the 553-Hz component and 753-Hz component. Once again, only one component was interaurally delayed during the middle 800 ms of each interval. In these conditions, the interaurally delayed component is never presented in isolation. Thresholds were measured for pulse durations of 50 and 100 ms in these conditions with complete temporal overlap of the pulses.

In a third set of conditions, the pulses temporally overlapped for all but 25 ms of each burst. The stream of 553-Hz pulses was shifted in time by the duration of one pulse minus 25 ms. (See the bottom portion of Figure 4.) As a result, during the 50-ms pulse condition, the 553-Hz and 753-Hz pulses overlapped for 25 ms; during the 100-ms pulse condition, they overlapped for 75 ms. These were the only two durations run in this set of conditions.

Additional thresholds were measured for pulse streams of each

Figure 4. A depiction of one interval each of the completely overlapping (top) and 25-ms non-overlap (bottom) stimuli presented in Experiment II. The cross-hatched portions represent 100-ms diotic pulses consisting of both the 553-Hz and 753-Hz components at the beginning and end of each interval. Solidly-filled portions of the figure, either white or black represent one component in isolation, as shown in the key. Striped portions of the figure indicate overlapping components, one with an interaural delay, one without. This figure roughly represents the 100-ms pulse condition with complete overlap of components and 25 ms of non-overlap.



553 Hz, LEFT
753 Hz, CENTER
2 COMPS, 1 DELAYED
2 COMPS, DIOTIC
100 ms
800 ms
100 ms
COMPLETE OVERLAP:
25-ms NON-OVERLAP:
TIME

frequency in isolation at pulse durations of 20, 50, and 100 ms. In these conditions, each interval was preceded and followed by a 100-ms diotic burst of the single frequency to remain consistent with the other conditions.

A three-down/one-up adaptive psychophysical procedure was used to estimate thresholds (Levitt, 1971). Following each incorrect response, the interaural delay was increased by a fixed amount (1 μs) for the following trial. After three consecutive correct responses, the interaural delay was decreased by 1 μs for the following trial. Because thresholds in many cases were very low (below 5 μs), a step size of 1 μs was necessary to prevent subjects from tracking down to an interaural delay of 0 μs during an experimental run. Each experimental run was made up of 70 trials. The first four reversals of each run were discarded and of the remaining reversals, the final even number of reversals were averaged to obtain an estimate of threshold IDT. An experimental run was discarded when fewer than ten total reversals occurred during the run. Six to eight threshold estimates were obtained in this manner for each experimental condition, with the median threshold estimate recorded as the final threshold for that condition.

The equipment used to generate and present stimuli and record responses was identical to that of Experiment I. The three subjects in this experiment were the same subjects who participated in Experiment I.

RESULTS II AND DISCUSSION II

Figure 5 shows the results for each of the three subjects, with each panel representing data from a single subject. Threshold IDT's are plotted against pulse duration. Open circles represent thresholds measured in the completely non-overlapping conditions, triangles represent the 553-Hz pulses in isolation, squares represent the 753-Hz pulses in isolation. Data from the 25 ms of non-overlap conditions are plotted as plus signs, and data from the complete overlap conditions are plotted as stars. The vertical lines through the stars show the interquartile ranges of the threshold estimates in the complete overlap conditions. Note that the ordinates of the three graphs are scaled differently.

For Subject 1, the interquartile ranges for the remaining thresholds were all less than 3 μs in size, with the exception of that for the 100-ms pulse duration in the 25 ms of non-overlap condition which was 3.83 μs in size. The remaining interquartile ranges for Subject 2 were less than 8 μs, except for that for the 553-Hz, 100-ms pulses in isolation, which had an interquartile range from 6.00 μs to 16.17 μs while the median threshold estimate was 14.07 μs. For Subject 3, the remaining interquartile ranges were all less than 4 μs in size, except for the 20-ms pulse duration, non-overlap condition (5.50 μs) and the 50-ms pulse duration, 25 ms of non-overlap condition (4.47 μs). In the following discussion of the results, a difference between thresholds will be considered significant when there is no overlap between the interquartile ranges for those thresholds.

The most obvious result is the difference between the complete overlap thresholds and those of all the other conditions. Consistent with the results of Experiment I, threshold IDT's are significantly larger when the pulses are presented without any isolated portions. The interference between components is not as large as in Experiment I because that experiment used seven-component complexes, where six components were distractors, while the present experiment used only two components, in effect, one target and one distractor in each interval.

Thresholds in the non-overlapping conditions (open circles) are quite similar across pulse durations, with a slight increase at the 20-ms pulse duration, particularly for Subject 2, for whom the 20-ms threshold is significantly different from the 50-ms threshold. This is probably due in part to confusion, because the pulses begin to lose the distinctive pitches of their carriers at these short pulse durations, so it becomes more difficult to attend to one pitch or the other.

Thresholds also begin to increase slightly, though non-significantly, at the longest pulse duration, 200 ms, in the non-overlapping pulse condition. This might reflect the fact that stream segregation is less likely to occur as the pulse duration increases, a common observation in streaming experiments (Bregman, 1990). In the present experiment, it is possible that the streams are no longer easily segregated on the basis of pitch at the 200-ms pulse duration. Perhaps confusion again results when the timing of the pulses promotes stream fusion but the apparent positions of the pulses promote stream segregation. In any case, the support for this idea is weak, given the non-significant increase in thresholds at the 200-ms pulse duration.

Thresholds for the conditions containing 25 ms of non-overlap (plus signs) are less than 5 μs larger than those in the completely non-overlapping conditions (circles) for each subject. This difference is not significant, except for Subject 3 for the 100-ms pulse durations. Once again, this is consistent with the results of Experiment I: a brief isolated presentation of each component eliminates almost all of the interference between components.

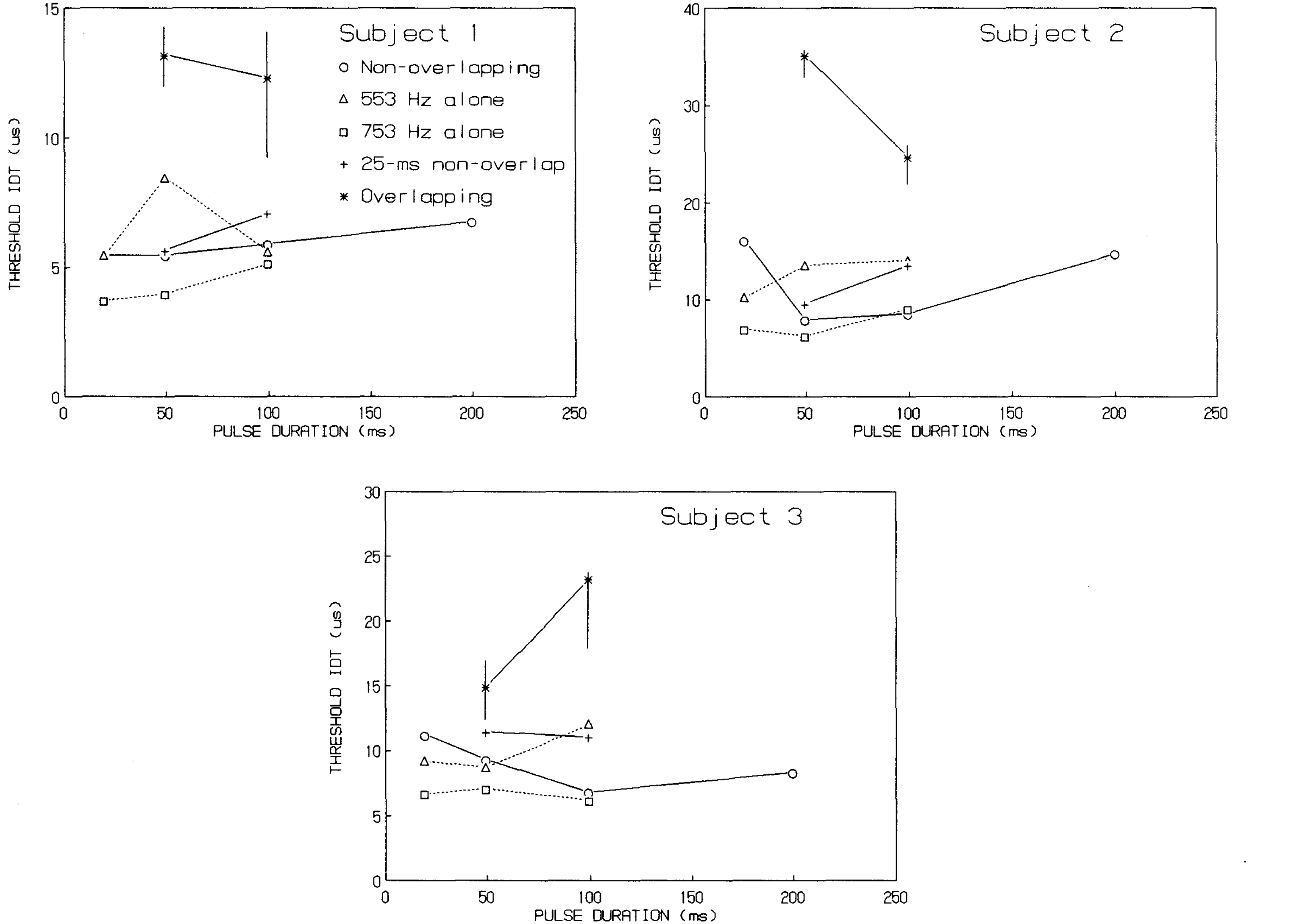
A small effect of frequency is observed. For all subjects, thresholds are significantly lower for the 753-Hz component in isolation (squares, dashed line) than for the 553-Hz component in isolation (triangles, dashed line) with the following exceptions: for a pulse duration of 100 ms for Subjects 1 and 2, and for a pulse duration of 50 ms for Subject 3.

The number of looks at each tone is apparently not of importance in the current experiment. Thresholds are fairly constant or increase as pulse duration decreases (number of pulses increases).

The most noteworthy result of the present experiment is that which was first identified: threshold IDT's are much higher when components to be lateralized completely overlap temporally with other spectral components than when the components to be lateralized are presented briefly in isolation. It seems very likely in the present experiment that the completely overlapping pulses are perceived as a single stream of complex events, so it is probably not surprising that interference between binaural information occurs across frequencies in those cases.

In terms of the "model" discussed in Experiment I, if a single descriptor is produced for the stream of completely overlapping pulse, a single description of its spatial position, based on a combination or average of the binaural information carried by the two frequencies, would be produced. It is significant that very brief isolated presentations of the components (25 ms in the present experiment) brings thresholds into the range of thresholds measured for the tones in isolation, even when substantial temporal overlap between components exists (25 ms and 75 ms of overlap for the 50-ms and 100-ms conditions, respectively). These results support the observations and conclusions of Experiment I and previous experiments by the author (Stellmack, 1990): differences between the temporal patterns of sinusoidal components may lead to segregation of pitches, but the additional factor of isolated presentation, albeit brief, is necessary for accurate spatial segregation.

Figure 5. Median estimates of threshold interaural differences of time (in μs) are plotted as a function of the pulse duration (in ms) for three subjects. Refer to the text for a complete description of the condition represented by each symbol. Vertical lines represent interquartile ranges for the complete overlap conditions. The ordinates are scaled differently in each panel.


Subject 1
o Non-overlapping
Δ 553 Hz alone
□ 753 Hz alone
+ 25-ms non-overlap
* Overlapping
THRESHOLD IDT (us)
PULSE DURATION (ms)
Subject 2
THRESHOLD IDT (us)
PULSE DURATION (ms)
Subject 3
THRESHOLD IDT (us)
PULSE DURATION (ms)

EXPERIMENT III

Auditory stimuli that naturally occur in the real world are normally more complex than the stimuli utilized in Experiments I and II. A naturally occurring sound source usually produces stimuli consisting of many spectral components whose frequency and/or amplitude are modulated over time in the same way across frequencies. Sounds emitted from different sources will almost certainly have different patterns of frequency- and amplitude-modulation. Experiment III utilizes stimuli that are more "realistic" than those in Experiments I and II in that the stimuli in the current experiment have such modulation.

A narrow band of noise contains fluctuations in amplitude that occur at a frequency related to the bandwidth of the noise: as the bandwidth of the noise increases, the fluctuations occur more rapidly. When a narrow band of noise is generated digitally by adding together a discrete series of sinusoids over a given frequency range, the specific pattern of amplitude fluctuations, or the stimulus <u>envelope</u>, is determined by the relative starting phases and amplitudes of the individual spectral components of the noiseband. Bands of noise with different center frequencies that are generated by adding together the same number of sinusoids will have identical, or coherent, envelopes if the same series of starting phases are used for the components of those noise bands. When the starting phases of the spectral components of two bands of noise are randomly and independently selected, the resulting bands of noise will have different, or incoherent, envelopes. Two bands of noise with coherent envelopes will overlap completely in time in terms of their amplitude fluctuations. On the other hand, two bands of noise with incoherent envelopes will vary over time such that during brief portions of their presentation, one noiseband will appear in relative isolation as the amplitude of the other becomes very small in the course of its random fluctuations.

Note that the terms "coherent" and "incoherent" are used in this experiment exclusively in reference to the relationship between the target and distractor envelopes. When the target and distractors have identical envelopes, the target and distractors will be described as coherent. When the target and distractors have different envelopes, the target and distractors will be described as incoherent. In the present experiment, the target and distractors are gated on and off simultaneously.

Trahiotis and Bernstein (1990) performed an experiment in which listeners attempted to lateralize a narrow band of noise in the presence of diotic noise distractors. The target had a center frequency of 500, 1000, 2000, 3000, or 4000 Hz and the target bandwidth was 40% of the center frequency (yielding a 200-Hz bandwidth at 500 Hz and 400-Hz bandwidth at 1000 Hz). The distractor consisted of band reject noise that immediately flanked the target band, such that the stimulus had a continuous spectrum up to 10,000 Hz. For low-frequency targets centered at 500 and 1000 Hz, which are most relevant to the present experiment, it appears that interference of a magnitude seen in previous experiments by the author involving pure tones occurs (Stellmack, 1990), although

Trahiotis and Bernstein describe the results as showing "vanishingly small amounts of interference" (Trahiotis and Bernstein, 1990, p. 812). Other experiments demonstrated little or no interference when listeners attempt to lateralize a low-frequency band of noise in the presence of diotic distractor bands (McFadden and Pasanen, 1976; Zurek, 1985).

These results are in apparent contradiction with the present series of experiments and previous studies by the author and colleagues (Stellmack, Dye, and Jakubczak, 1989; Dye, 1990) in which large amounts of binaural interference are observed across frequencies. The fact that experiments in which little or no binaural interference was observed across frequencies used bands of noise while experiments in which binaural interference was observed used pure tones results. There are two important aspects of the manner in which stimuli were generated in the noise band experiments that might account for the differences between the results of noise band and pure tone experiments: 1) the fact that the stimuli used in the noise band experiments were generated randomly and independently, and 2) the noise band targets and distractors had different bandwidths from one another. Both of these factors produce targets and distractors with incoherent envelopes.

The results of Experiments I and II suggest the possibility that a narrow-band noise target with an interaural delay will be more readily lateralized in the presence of an incoherent distractor than with a coherent distractor. The brief intervals of relative isolation of the noisebands that occur when the envelopes are incoherent might be sufficient to allow listeners to detect the interaural delay of the target band, as in Experiment I. If the target and distractor envelopes are coherent, the target and distractor will completely overlap in time, which would be expected to result in relatively large amounts of binaural interference between the target and distractor.

In order to test these predictions, the narrow-band noise stimuli cannot be generated randomly. To study the effects of the relationship between specific target and distractor envelopes, one must record or save each series of starting phases, if not the actual stimulus, so that the envelopes of interest can be reproduced in different experimental trials at different center frequencies. When certain noise samples are repeatedly presented during a block of trials, they are described as reproducible or frozen noise samples.

The following experiment makes use of reproducible noise to study envelope-dependent effects on the ability of listeners to lateralize a target noise band in the presence of a distractor noise band. The use of reproducible noise samples will allow for assessment of the relationship between target and distractor envelopes and lateralization performance.

METHODS III

Three 500-ms intervals were presented during each experimental trial, with 300 ms of silence between intervals. The first interval, the cue, always consisted of a diotic presentation of the target band in isolation and was intended to allow subjects to attend to the pitch of the target, if possible. The remaining two intervals, the listening intervals, consisted of three noise bands: the target and two flanking distractor bands. The target band was interaurally delayed in one of the listening intervals and diotic in the other. The distractors were always presented diotically. Subjects were instructed to identify the interval in which the target was interaurally delayed.

The target and distractors were of equal bandwidth within a block of trials. Bandwidths of 10 and 20 Hz were studied. In the 10-Hz bandwidth conditions, the target band was centered at 750 Hz with the distractor bands centered at 650 and 850 Hz. In the 20-Hz bandwidth conditions, the target band was again centered at 750 Hz and the distractor band center frequencies were either 50, 100, or 200 Hz above and below the center frequency of the target. Linear rise/decay times of 50 ms were used to gate the target and distractor on and off.

Stimuli were generated by performing a 32,768 point inverse Fast Fourier Transform (FFT) on a Masscomp minicomputer at a sampling rate of 20,000 Hz, providing a resolution of .61 Hz between spectral components. Only the first 10,000 time-domain points (500 ms) were used. When stimuli are generated in this way, amplitudes and phases are provided for each discrete frequency within the desired band and the inverse FFT produces the corresponding digital time-domain waveform. In the present experiment, all components had equal amplitudes. For each bandwidth, sets of starting phases were randomly generated and stored on computer. (The starting phases are shown in Appendix A.) When amplitudes are randomly selected for the individual spectral components from a Rayleigh distribution, the average number of envelope peaks per second is given by .6411 times the bandwidth of the noise (Rice, 1954). Equal-amplitude noise, as used in this experiment, is virtually indistinguishable from Rayleigh-distributed noise when more than 12 spectral components are present (Hartmann, 1987). In this experiment, 17 components were used to generate the 10-Hz wide noisebands and 34 components were used to generate the 20-Hz wide noisebands, so the formula for computing the average number of envelope peaks per second can be applied in this case. Given that formula and the fact that the listening intervals were 500 ms in duration, an average of 3-4 peaks and 6-7 peaks could be expected in the 10-Hz bandwidth and 20-Hz bandwidth stimuli, respectively.

In each block of trials, three of the sets of starting phases were used to generate three target bands and three distractors consisting of the two distractor bands with identical envelopes. For the three target bands, two different versions of the target were generated, one diotic and one containing an interaural delay. In the delayed version of the target, the starting phases in the left channel were advanced such that the interaural difference of time was equal for the components of the target band. Thus, the target would be lateralized to the left if the target were presented in isolation. When the components in a band of noise have equal interaural differences of time, this produces a delay of the whole waveform, both the envelope and fine structure (Henning, 1980).

On each experimental trial, one of the three targets and one of the three distractors were randomly selected and presented in both listening intervals, with the target interaurally delayed in one interval. As a result, there were nine different combinations of targets and distractors that could be presented during each trial, three of which consisted of a target and distractor with identical envelopes. Responses were accumulated separately and d′ was computed for each of the nine possible stimuli in a block of trials.

A block of trials consisted of 100 trials, in two groups of 50. Before each group of 50 trials, the subject was allowed to listen and respond to practice trials, for which responses were not recorded. When ready, subjects initiated each set of 50 experimental trials by pressing a button on the response terminal. Nine blocks of 100 trials were run for each group of three stimulus envelopes, so that an average of 100 trials were presented for each target-distractor pair. The target IDT was constant across the nine blocks of trials for a given condition, and was selected to yield 80-85% correct in a block of 100 trials as determined by each individual subject's performance. As a result, the target IDT's were not equal across subjects in each condition. Data were gathered during two-hour sessions during which subjects were run individually. A typical experimental session consisted of 500-600 trials per subject.

The equipment used to generate and present stimuli was identical to that in Experiments I and II with the following exception: stimuli were presented to Subjects 2 and 4 through Telephonics TDH-49 earphones suspended in Auraldomes, while stimuli were presented to Subjects 1 and 3 through Sony MDR-V6 headphones.

Subject 3 in the present experiment was also Subject 3 in Experiments I and II. Subject 1 was an undergraduate volunteer from the author's university who had extensive experience from participation in some of the author's previous lateralization studies. Subjects 2 and 4 were also undergraduate volunteers from the author's university, but with no previous experience in lateralization experiments. Subjects 2 and 4 listened to stimuli similar to those presented in the experiment over the course of about two weeks prior to actual data collection as training for the experiment. Subject 4 provided data for only two of the 20-Hz bandwidth conditions and then dropped out of the experiment due to illness. All subjects were paid an hourly wage for their participation in the experiment.

RESULTS III

The results for a bandwidth of 20 Hz and a frequency spacing of 100 Hz between center frequencies are shown in Figure 6. ("Frequency spacing" and "Δf" in the following discussion refer to the frequency spacing between the center frequencies of the target and distractor bands.) The bandwidth, center frequencies, and codes for identifying the envelopes used in the conditions represented in the figure are at the top of the figure. Each envelope is identified by a 3-digit number in which the first two digits indicate the bandwidth and the third digit is an arbitrary identifier. In each panel, d′ is plotted as a function of the target-distractor pair. The numbers 1, 2, and 3 for the targets and distractors in each panel correspond to the three envelopes listed at the top of the figure. In this figure, for example, Target 1 and Distractor 1 both have Envelope 204, Target 2 and Distractor 2 have Envelope 205, and Target 3 and Distractor 3 have Envelope 206. Each panel in the figure represents data from a different subject. The ordinates are scaled to accommodate each subject's data. The interaural difference of time (IDT) at which the target was presented to each subject is shown in each panel.

In conditions in which the target and distractor have the same envelope, for example Target 1 and Distractor 1, the target and distractors are coherent, as described earlier. The target and distractors are incoherent in conditions in which they have different envelopes, for example Target 1 and Distractor 2 or Target 1 and Distractor 3.

As predicted, performance was poorest when the target and distractors had identical envelopes, with one exception. For each distractor, d' is smallest for the target with the same envelope except for Subject 2, Distractor 3, for which Target 1 produced poorest performance.

Looking at the data with respect to each target, in most cases, performance was worst when the distractor is coherent (i.e., has the same envelope). For example, for Target 3 (squares), d' is smaller for all subjects with Distractor 3 than with either Distractor 1 or Distractor 2. The same is true for the other targets with one exception: for Subject 3 and Target 1, Distractor 2 produced slightly poorer performance than Distractor 1.

The results of all of the 20-Hz bandwidth conditions run in this experiment are shown in Figures 7 and 8 in the same form as Figure 6, with the data from Figure 6 shown again in Figure 7a. Figure 7b shows results for a frequency spacing of 100 Hz, as in Figure 7a, but with different envelopes (201, 202, and 203). Figures 8a and 8b show data for frequency spacings of 50 and 200 Hz, respectively, with the same set of envelopes as in Figure 7a (204, 205, and 206). The similar pattern of results across subjects in each condition is striking given that they had very different sensitivities to IDT's of the target.

The data for Envelopes 201, 202, and 203 (Figure 7b) are similar to those for Envelopes 204, 205, and 206 (Figure 7a and Figure 6). For a given distractor, performance was poorest when the target was coherent with two exceptions: for Subject 2, Target 2 produced slightly poorer performance than Target 1 when paired with Distractor 1, and for Subject 4, Target 3 produced slightly poorer performance than Target 1 with Distractor 1. In addition, for a given target, poorest performance occurred when it was paired with the distractor that had an identical envelope, again with one exception (Subject 2, Target 2).

With only the few exceptions noted, the data for the 20-Hz bandwidth and $\Delta f = 100$ Hz support the prediction that performance would be poorest when the target and distractors are coherent.

The results for the 20-Hz bandwidth and $\Delta f = 50$ Hz (Figure 8a) are similar to those for $\Delta f = 100$ Hz (Figure 7a) in that for each distractor, poorest performance occurred when the target and distractor were coherent. In addition, the poorest performance for each target was observed when it was presented with the distractor having the same envelope.

For Subjects 2 and 3, when $\Delta f = 50$ Hz, performance was much poorer in the coherent conditions relative to the incoherent conditions than when $\Delta f = 100$ Hz. For Subject 2, d′ actually becomes slightly negative in two of the coherent conditions. A negative d′ suggests that the subject was responding to the apparent movement of the stimulus across intervals and the stimulus with the delayed target was lateralized toward the opposite side of the head, although in this case d′ is very close to zero, so the negative d′ may result from random variation about chance performance. The pattern of results for Subject 1 is identical for $\Delta f = 50$ Hz and $\Delta f = 100$ Hz.

The effects of frequency spacing can be examined, since this is the only difference between conditions represented in Figures 7a, 8a, and 8b. As frequency spacing increased from 50 to 100 to 200 Hz, the interference between coherent targets and distractors was reduced. At Δf = 200 Hz (Figure 8b), performance was almost exclusively dependent upon which target was presented, independent of the distractor with which it was paired. For example, Target 3 produced better performance than Targets 1 and 2, except for Subject 3, Distractor 3 where performance was nearly equal with all three targets.

The reduction of the target-distractor interaction can be more clearly seen in Figure 9, which displays the 20-Hz bandwidth data in bar chart form. T1, T2, and T3 denote Targets 1, 2, and 3 respectively. The solid bars show the mean d′ of all of the subjects in each condition. The difference between the means of the coherent and incoherent conditions is largest for Δf = 50 Hz (Figure 9a), but becomes smaller at Δf = 100 Hz (Figure 9b), with no consistent differences between the means independent of distractor for Δf = 200 Hz (Figure 9c).

The data for the 10-Hz bandwidth and Δf = 100 Hz conditions are shown in line chart form in Figure 10 and bar chart form in Figure 11. Performance was not consistently poorer in the coherent conditions relative to the incoherent conditions as it was in the 20-Hz bandwidth conditions for Δf = 50 Hz and 100 Hz. However, envelope-specific effects can still be observed. Note that when a target with one envelope produced consistently high or low performance across distractors, distractors with that same envelope produced the opposite effect. For example, in Figure 10a, for Envelopes 101, 102, and 103, poorest performance was seen with Target 2 in most cases, while

Distractor 2 generally produced better performance. Similarly, in Figure 10b, for Envelopes 104, 105, and 106, performance was poorest for Target 1 across all distractors, but performance was relatively good with Distractor 1, particularly for Subject 2. It is as if a particular envelope causes the noise band to dominate the stimlus in terms of spatial information, resulting in poor performance when the target has that envelope and good performance when the distractor has that envelope, or vice versa.

The prediction that target bands of noise would be more difficult to lateralize in the presence of coherent versus incoherent distractors was confirmed in only some of the conditions studied in this experiment. However, envelope-specific effects were identified in all conditions, as were some effects of frequency spacing.

Figure 6. The values of d′ are plotted for each of the target-distractor pairs for four subjects. Targets and distractors had a bandwidth of 20 Hz. The center frequency of the target was 750 Hz and the center frequencies of the distractors were 650 and 850 Hz. The labels 1, 2, and 3 for the targets and distractors in each panel correspond to Envelopes 204, 205, and 206, respectively, as indicated at the top of the figure. The ordinates are not scaled identically.

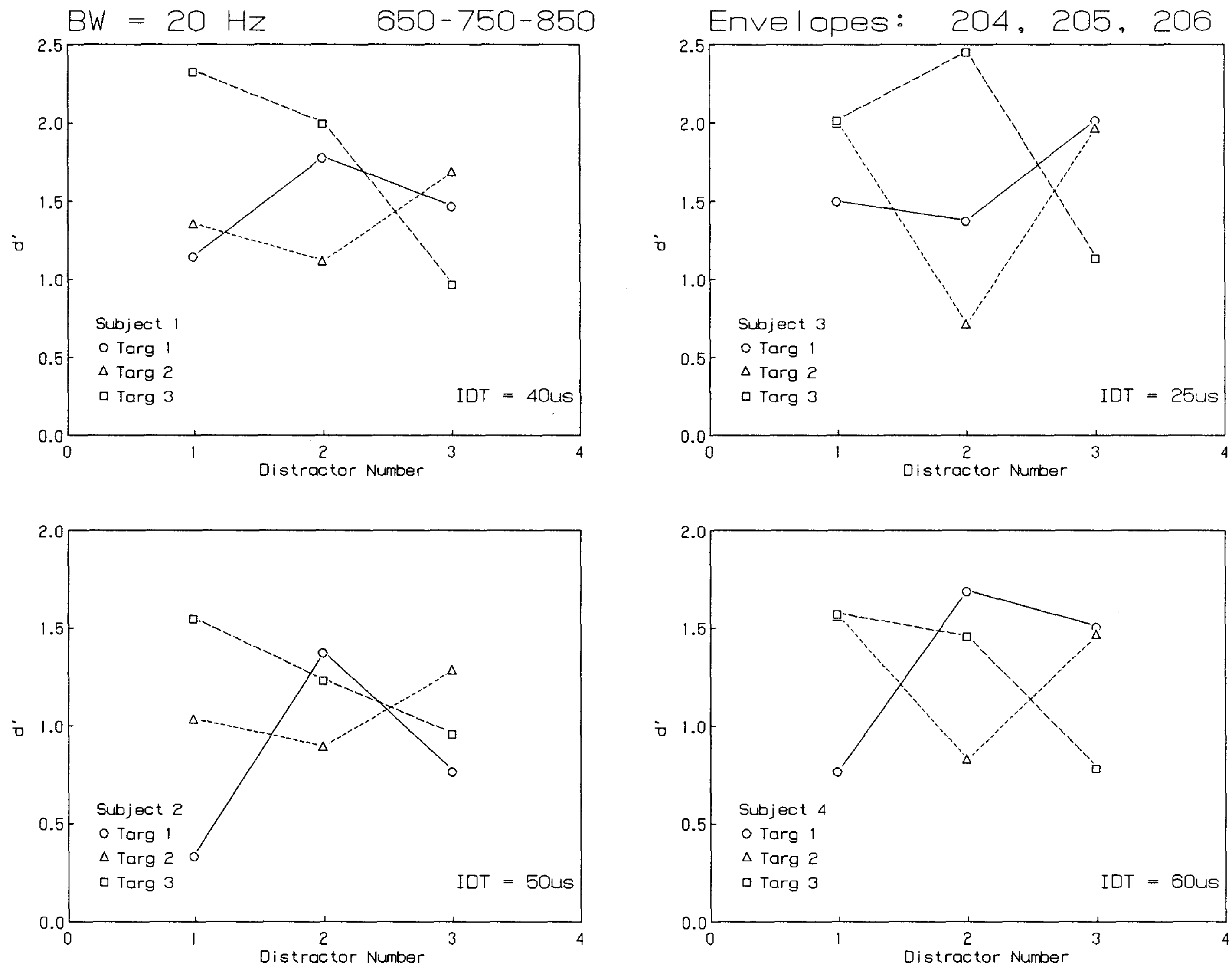
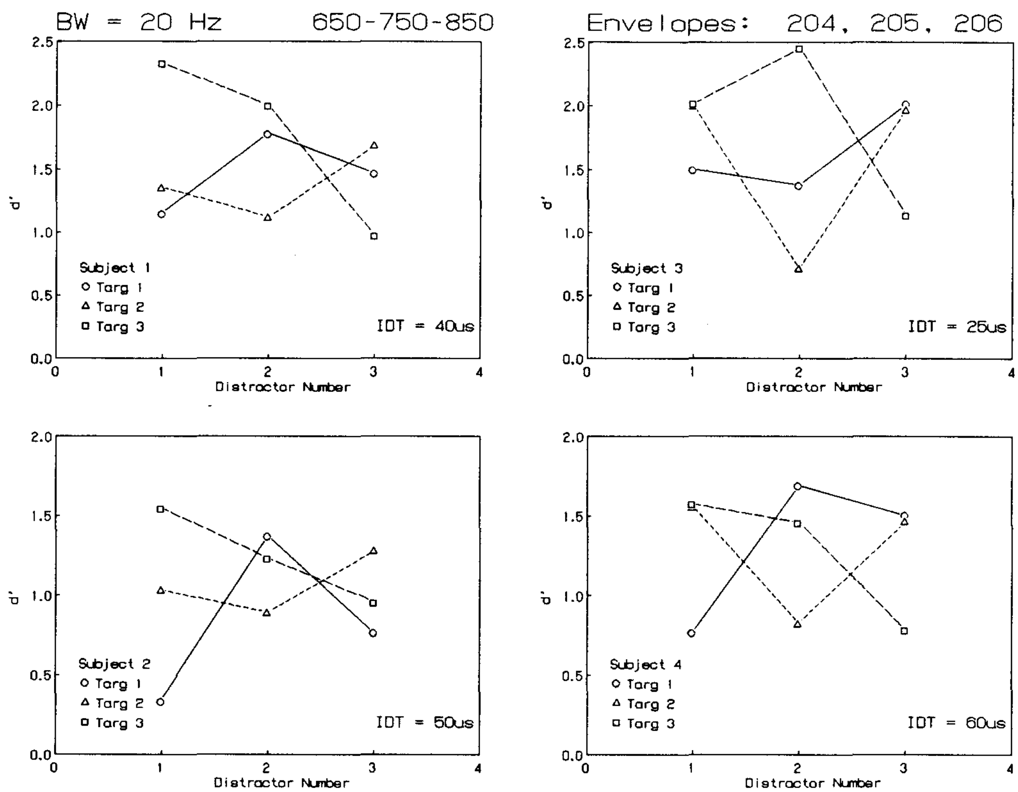
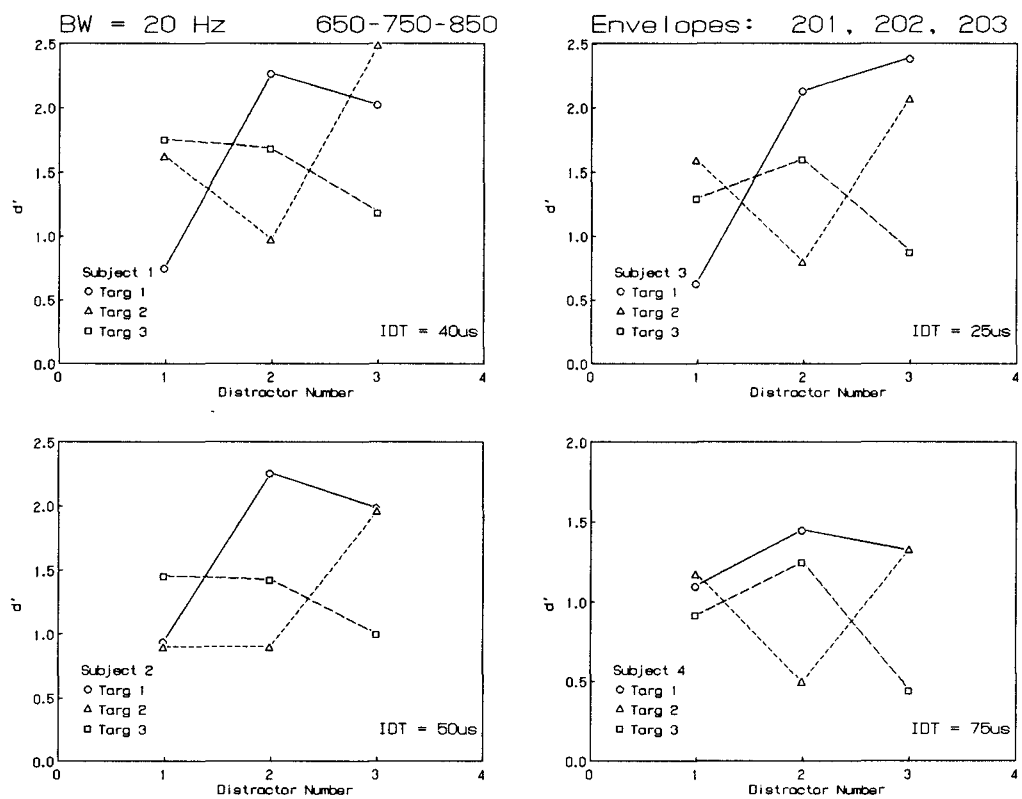

BW = 20 Hz 650-750-850
Envelopes: 204, 205, 206
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
Targ 1
Targ 2
Targ 3
IDT = 25us
Subject 2
Targ 1
Targ 2
Targ 3
IDT = 50us
Subject 4
Targ 1
Targ 2
Targ 3
IDT = 60us

Figure 7. The data for all of the 20-Hz bandwidth, $\Delta f = 100$ Hz conditions are shown in the same form as Figure 6. Figure 7a shows the data for envelopes 204, 205, and 206 (from Figure 6). In Figure 7b, data for $\Delta f = 100$ Hz with a different set of envelopes (201, 202, and 203) are shown.

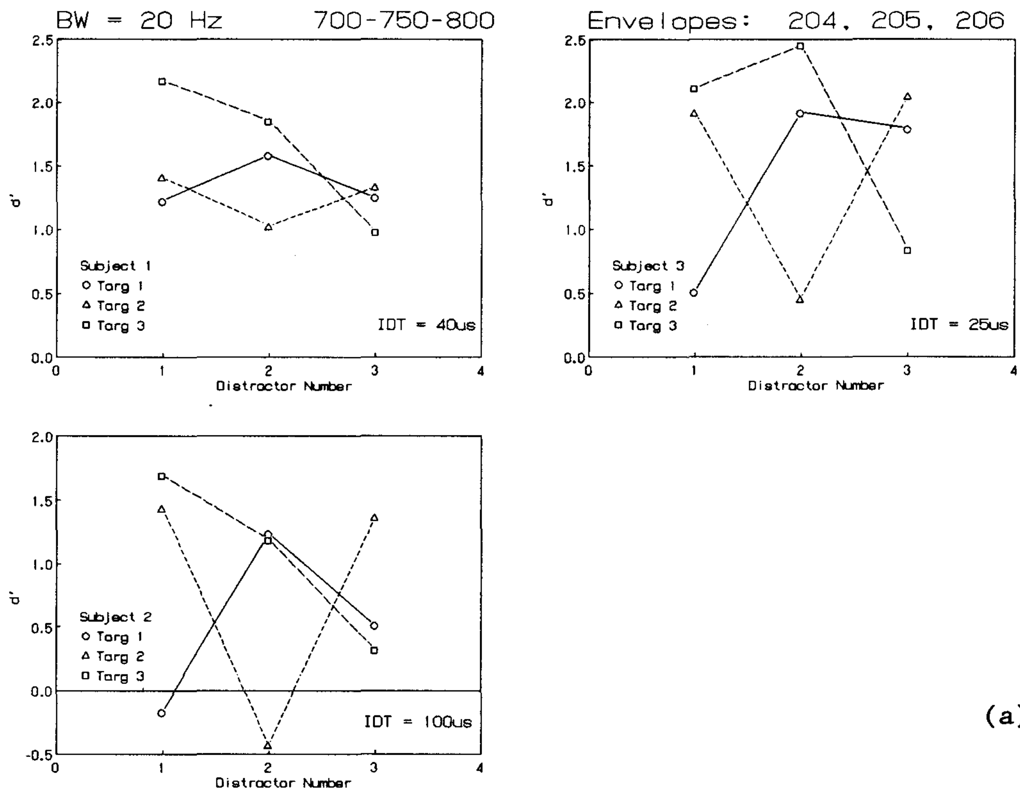

BW = 20 Hz 650-750-850
Envelopes: 204, 205, 206
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
IDT = 25us
Subject 2
IDT = 50us
Subject 4
IDT = 60us

(a)

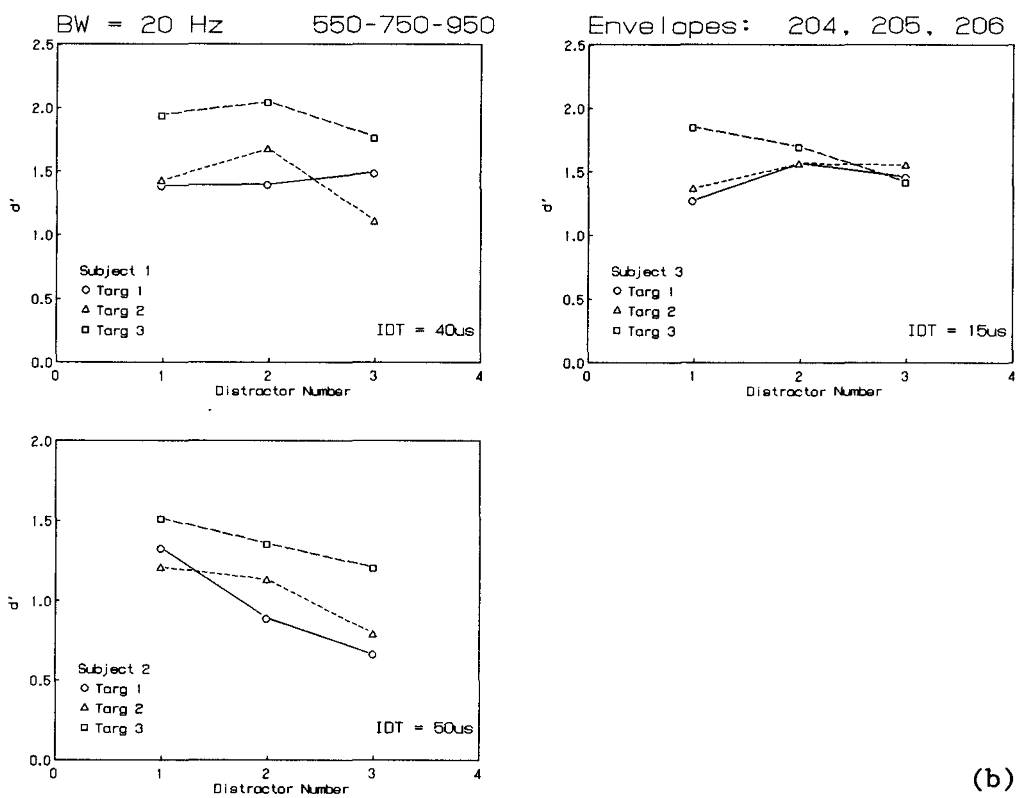

BW = 20 Hz 650-750-850
Envelopes: 201, 202, 203
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
IDT = 25us
Subject 2
IDT = 50us
Subject 4
IDT = 75us

(b)

Figure 8. The data for the 20-Hz bandwidth conditions with $\Delta f = 50$ Hz (Figure 8a) and $\Delta f = 200$ Hz (Figure 8b) are shown in the same form as Figure 6. In both Figures 8a and 8b, the same set of envelopes (204, 205, and 206) were used.

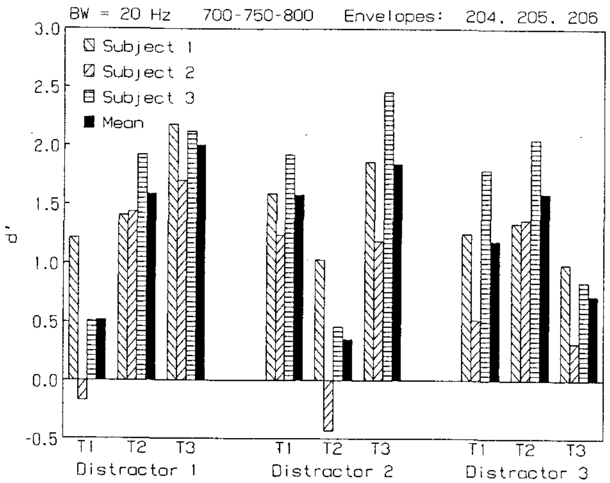

BW = 20 Hz 700-750-800
Envelopes: 204, 205, 206
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
Targ 1
Targ 2
Targ 3
IDT = 25us
Subject 2
Targ 1
Targ 2
Targ 3
IDT = 100us

(a)

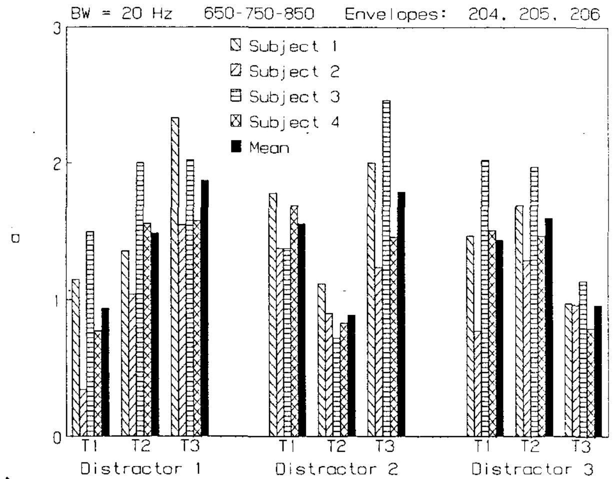

BW = 20 Hz 550-750-950
Envelopes: 204, 205, 206
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
Targ 1
Targ 2
Targ 3
IDT = 15us
Subject 2
Targ 1
Targ 2
Targ 3
IDT = 50us

(b)

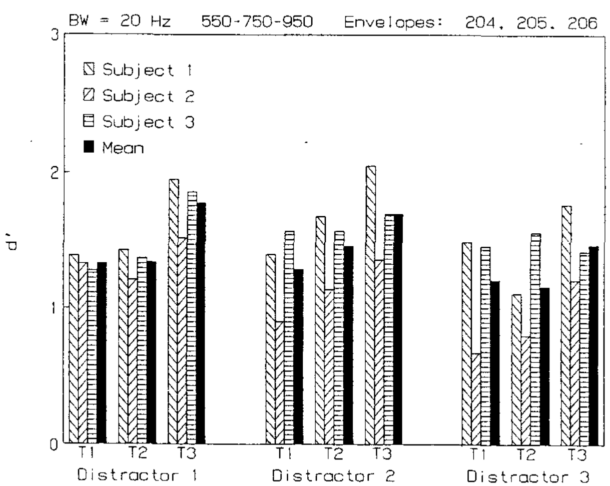
Figure 9. The data for all of the 20-Hz bandwidth conditions (from Figures 7 and 8) are shown as bar charts. Each bar represents the value of d' for a different subject, with the mean value of d' shown as a solid bar for each condition. The left-hand side of this figure corresponds to the conditions shown in Figure 8, and the right-hand side corresponds to the conditions of Figure 7.


BW = 20 Hz 700-750-800 Envelopes: 204, 205, 206
Subject 1
Subject 2
Subject 3
Mean
d'
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
T1 T2 T3
Distractor 1
T1 T2 T3
Distractor 2
T1 T2 T3
Distractor 3

(a)


BW = 20 Hz 650-750-850 Envelopes: 204, 205, 206
Subject 1
Subject 2
Subject 3
Subject 4
Mean
3
2
1
0
T1 T2 T3
Distractor 1
T1 T2 T3
Distractor 2
T1 T2 T3
Distractor 3

(b)


BW = 20 Hz 550-750-950 Envelopes: 204, 205, 206
Subject 1
Subject 2
Subject 3
Mean
d'
3
2
1
0
T1 T2 T3
Distractor 1
T1 T2 T3
Distractor 2
T1 T2 T3
Distractor 3

(c)

(d)

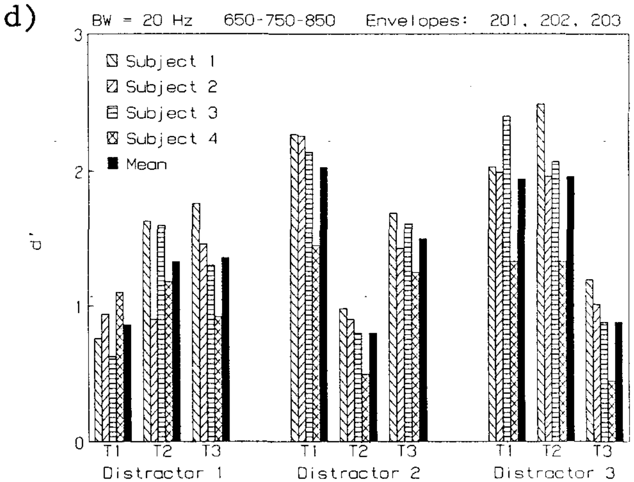
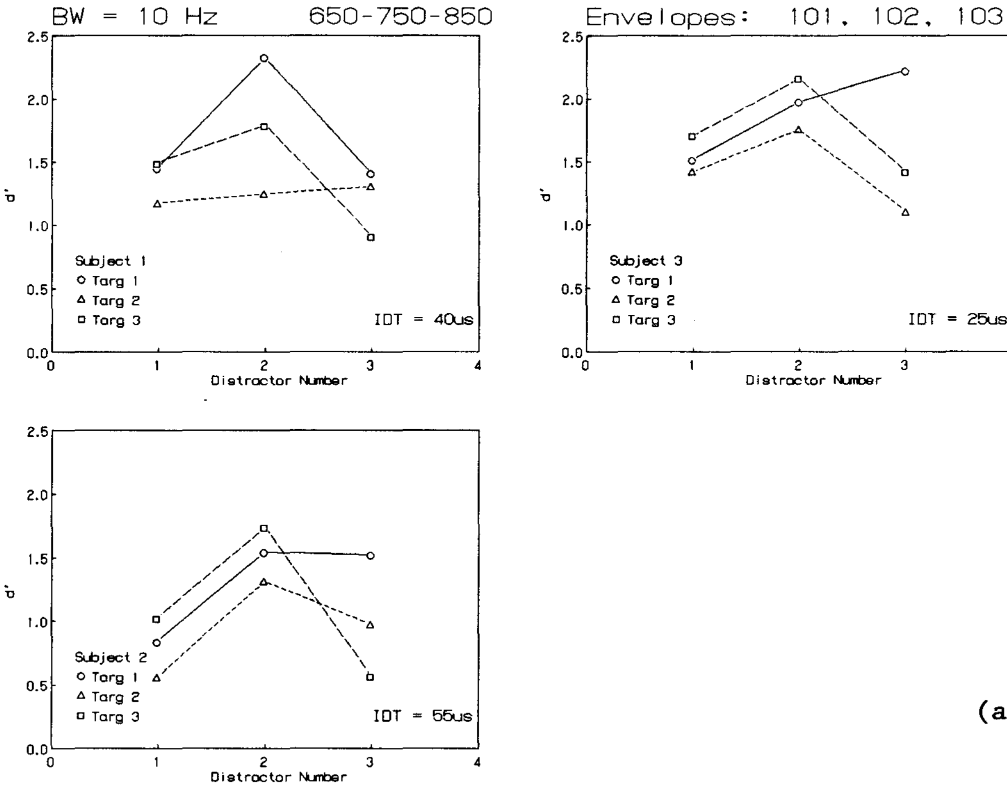
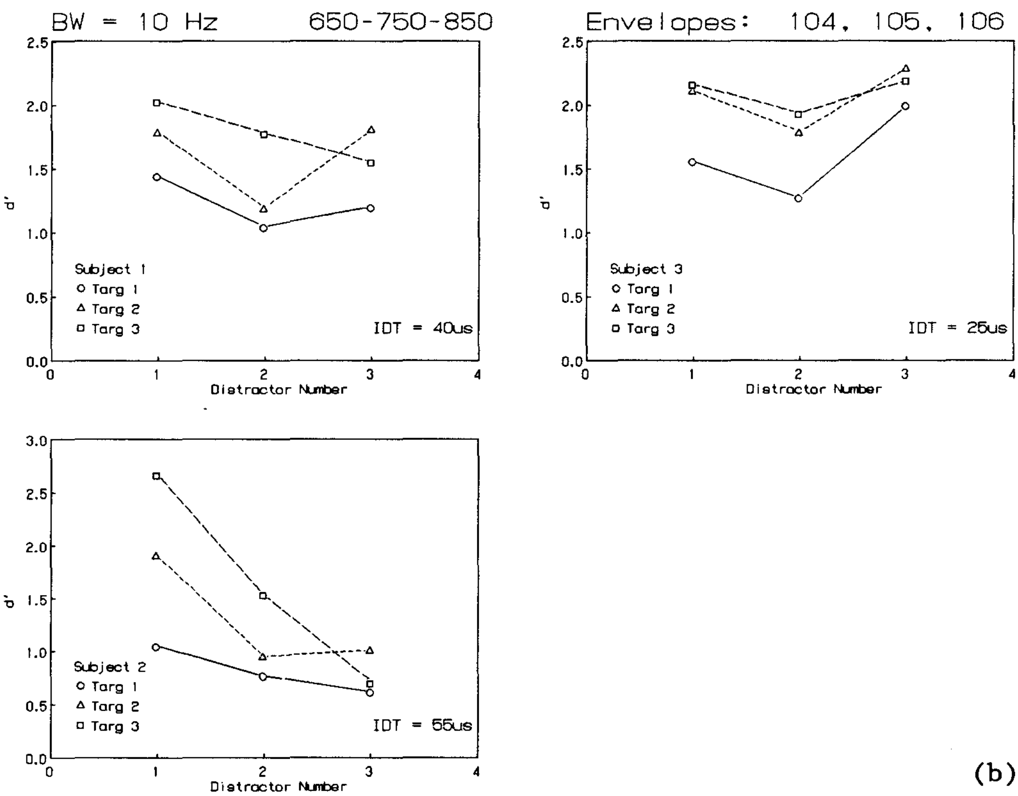

BW = 20 Hz 650-750-850 Envelopes: 201, 202, 203
Subject 1
Subject 2
Subject 3
Subject 4
Mean
d'
3
2
1
0
T1 T2 T3
Distractor 1
T1 T2 T3
Distractor 2
T1 T2 T3
Distractor 3

Figure 10. The data for all of the conditions in which the bandwidth was 10 Hz and $\Delta f = 100$ Hz are shown in the line chart form of Figures 7 and 8. Figure 10a shows data for Envelopes 101, 102, and 103. Figure 10b shows data for a different set of envelopes, 104, 105, and 106.


BW = 10 Hz
650-750-850
Envelopes: 101, 102, 103
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
IDT = 25us
Subject 2
IDT = 55us

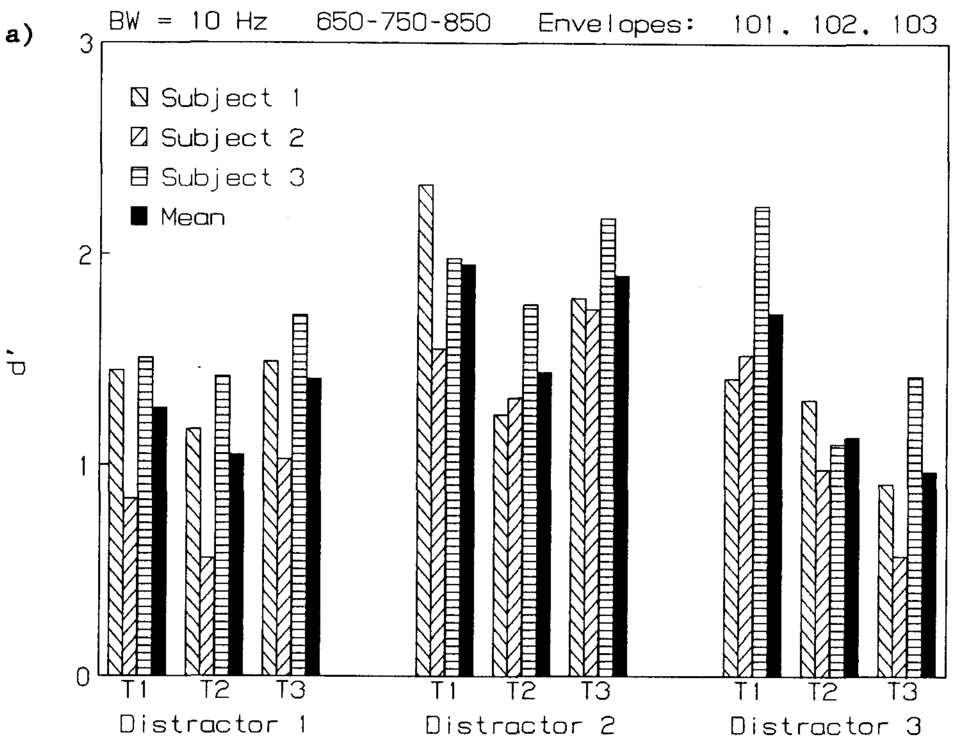
(a)


BW = 10 Hz
650-750-850
Envelopes: 104, 105, 106
d'
Distractor Number
Subject 1
Targ 1
Targ 2
Targ 3
IDT = 40us
Subject 3
IDT = 25us
Subject 2
IDT = 55us

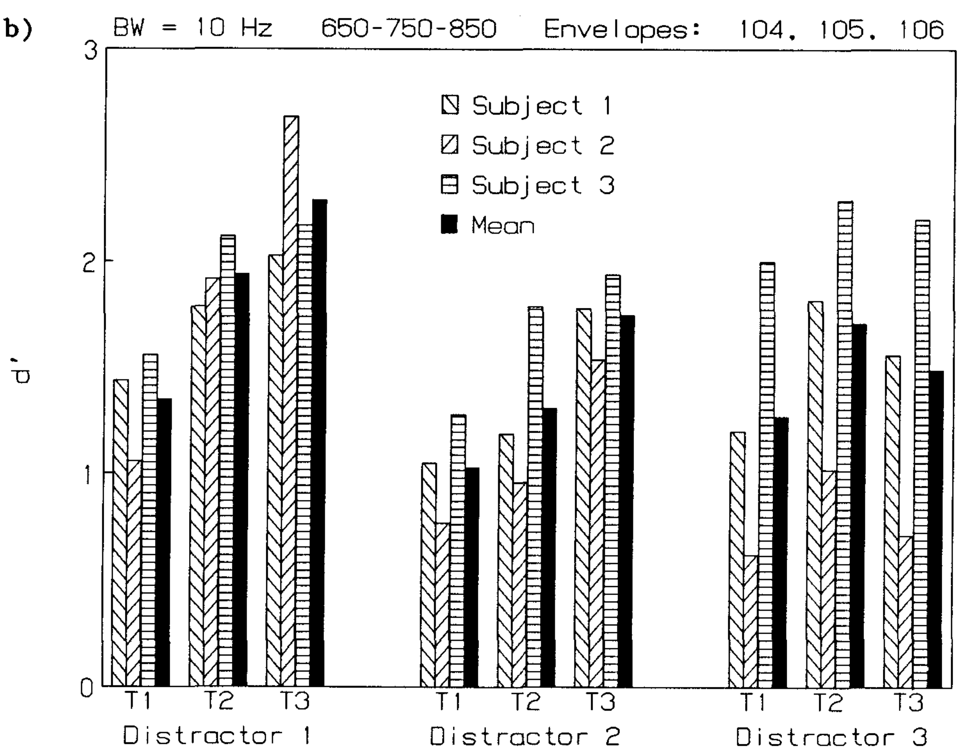
(b)

Figure 11. The data for all of the conditions in which the bandwidth was 10 Hz and $\Delta f = 100$ Hz (replotted from the previous figure) are shown in the bar chart form of Figures 9. Figure 11a shows data for Envelopes 101, 102, and 103 (from Figure 10a), and Figure 11b shows data for Envelopes 104, 105, and 106 (from Figure 10b).

a)


BW = 10 Hz 650-750-850 Envelopes: 101, 102, 103
Subject 1
Subject 2
Subject 3
Mean
d'
0
1
2
3
T1 T2 T3
Distractor 1
T1 T2 T3
Distractor 2
T1 T2 T3
Distractor 3

b)


BW = 10 Hz 650-750-850 Envelopes: 104, 105, 106
Subject 1
Subject 2
Subject 3
Mean
d'
0
1
2
3
T1 T2 T3
Distractor 1
T1 T2 T3
Distractor 2
T1 T2 T3
Distractor 3

DISCUSSION III

In this experiment, target noisebands were more difficult to lateralize in the presence of coherent distractors than in the presence of incoherent distractors when the bandwidth was 20 Hz and $\Delta f = 100$ Hz or less. This was not true for a bandwidth of 20 Hz and $\Delta f = 200$ Hz or for a bandwidth of 10 Hz and $\Delta f = 100$ Hz. However, in these cases, performance was still dependent on the particular envelopes of the target and distractor. All of these effects are undetectable when stimuli are generated randomly. Only when the stimuli are reproducible noise samples can such envelope-specific effects become evident. The robustness of these envelope-specific effects is further emphasized by the fact that similar patterns of results were obtained for subjects with very different sensitivities to IDT's of the target.

In order to identify some of the properties of the stimulus envelopes that may be responsible for the observed effects, it is necessary to examine the time-domain waveforms of the actual targets and distractors used in the present experiment. Figure 12 shows the six targets (bands of noise centered at 750 Hz) used in the 20-Hz bandwidth conditions in the present experiment. Figure 13 shows the six targets used in the 10-Hz bandwidth conditions. The distractors are not shown because the envelopes are identical to those of the targets, though the distractors had twice the peak amplitude of the targets because the distractors consisted of two noisebands added together.

It can be seen that the temporal fluctuations occur at a higher frequency in the 20-Hz wide noisebands than in the 10-Hz wide noisebands. In the 20-Hz wide noisebands, 5 or 6 major peaks occur during the 500-ms listening interval, while 3 or fewer large peaks occur in the 10-Hz wide noisebands, as predicted in the discussion of the stimulus generation in the Methods section. The number of abrupt onsets and offsets is correspondingly higher in the 20-Hz wide noisebands as a result.

One prediction might be that lateralization of the target improves as the correlation between the target and distractor envelopes diminishes, in other words, greater incoherence exists. This explanation cannot be correct. The correlation is equal between two different envelopes regardless of which envelope was given to the target and distractor. For example, the correlation between a target with Envelope 204 and a distractor with Envelope 206 is equal to the correlation between a target with Envelope 206 and a distractor with Envelope 204. This would predict identical performance in these two conditions if performance was a simple function of the correlation between envelopes. Since this is not the case (see, for example, Figure 8a), the correlation between the target and distractor envelopes cannot by itself account for the results.

In the 20-Hz bandwidth conditions, the coherent or incoherent relationship between target and distractor was able to account for the relative performance for $\Delta f = 50$ and 100 Hz. It is possible that the incoherent relationship between the target and distractor envelopes provides enough of an isolated presentation of the target to allow lateralization of the target, as in Experiments I and II. This brief isolated presentation would be of particular importance at the narrow frequency spacings in which portions of the target and distractor bands begin to interact within a monaural critical band, resulting in the interaction of binaural information as well. (The equivalent rectangular bandwidth of the monaural critical band at 750 Hz is about 100 Hz, from Moore and Glasberg, 1983.)

For $\Delta f = 200$ Hz, it was observed that performance was a function of the target envelope, independent of the distractor. For example, best performance generally resulted when Target 3, with Envelope 206, was presented compared to Targets 1 and 2, with Envelopes 204 and 205 (see Figure 8b). Envelope 206 has a smaller peak amplitude and appears to have less variability in its envelope fluctuations than both Envelopes 204 and 205. In other words, Envelope 206 is more constant in terms of its intensity during the duration of the listening interval. When presented with a distractor that has the widely fluctuating intensity of Envelope 204 or 205, the spatial information in a target with Envelope 206 might be weighted more heavily because of the more constant nature of the target. Similarly, performance was better for a target with Envelope 201 relative to performance for targets with Envelope 202 or 203. Once again, Envelope 201 appears to have a smaller peak amplitude and less envelope variability than Envelopes 202 and 203.

The idea of an inverse relationship between the variability of an envelope's fluctuations and the relative ability of listeners to lateralize a target with that envelope is also supported by the data for the 10-Hz bandwidth shown in Figure 10a. Generally, poorer performance was obtained for a target with Envelope 102. In examining the stimulus waveforms (Figure 13), Envelope 102 has a larger peak amplitude and appears to have greater variability in the amplitude of the envelope. This relationship does not hold for the data shown in Figure 9c. A target with Envelope 104 consistently results in poorer performance relative to Envelopes 105 and 106, but Envelope 104 clearly has lower variability than Envelopes 105 and 106. Perhaps there is some moderate level of envelope fluctuation that is optimal for the spatial information in a band of noise to dominate a stimulus, such that, if the fluctuations are either too great or too small, as in these examples, listeners are less able to lateralize the noiseband in the presence of additional bands of noise with different envelopes.

Recall that the interaural delay of the target was of both the envelope and fine structure. If the interaural delays of the envelope transitions that occur during the course of the stimulus were critical, it would be expected that envelopes with larger envelope variability would be lateralized more effectively because of the more sharply defined fluctuations of the envelope. Once again, this is not the case, leading to the conclusion that abrupt onsets and offsets are not as important in lateralizing the target as some other property of the noiseband envelope. Recall that, in Experiment I, when the target had no abrupt onsets or offsets during the listening interval, but the distractor did in order to produce a temporal notch, lateralization of the target approached that for the target in isolation.

This discussion of the relationships between the target and distractor envelopes has proceeded in a qualitative rather than quantitative fashion. In order to produce more definite conclusions as to the properties of the target and distractor envelopes that drive performance, it is necessary to examine many more envelopes in a larger number of bandwidth and frequency spacing conditions than were studied here. This experiment provides some good clues as to what properties of the envelopes might be important, and, at the very least, shows the importance of examining the stimuli in terms of specific envelopes of frozen samples of noise. This method allows an examination of the effects on lateralization of specific relationships between the target and distractor envelopes that is not possible when stimuli are generated randomly.

Figure 12. The time-domain waveforms of the target stimuli used in the 20-Hz bandwidth conditions are shown. Each stimulus is 500 ms in duration with 50-ms linear onsets and offsets. The amplitudes of the waveforms are scaled identically, and shown as the output voltage of the digital-to-analog converters.

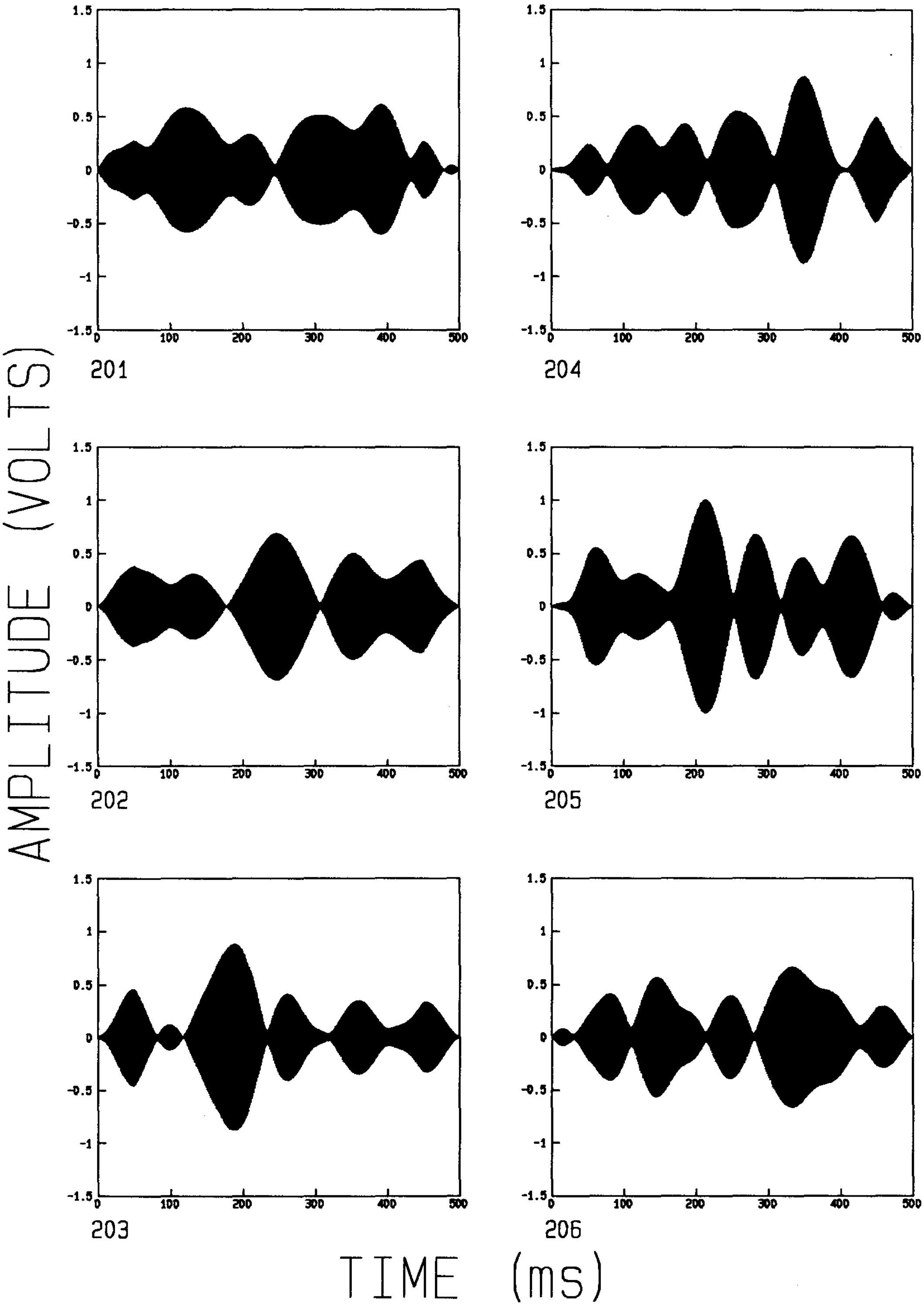
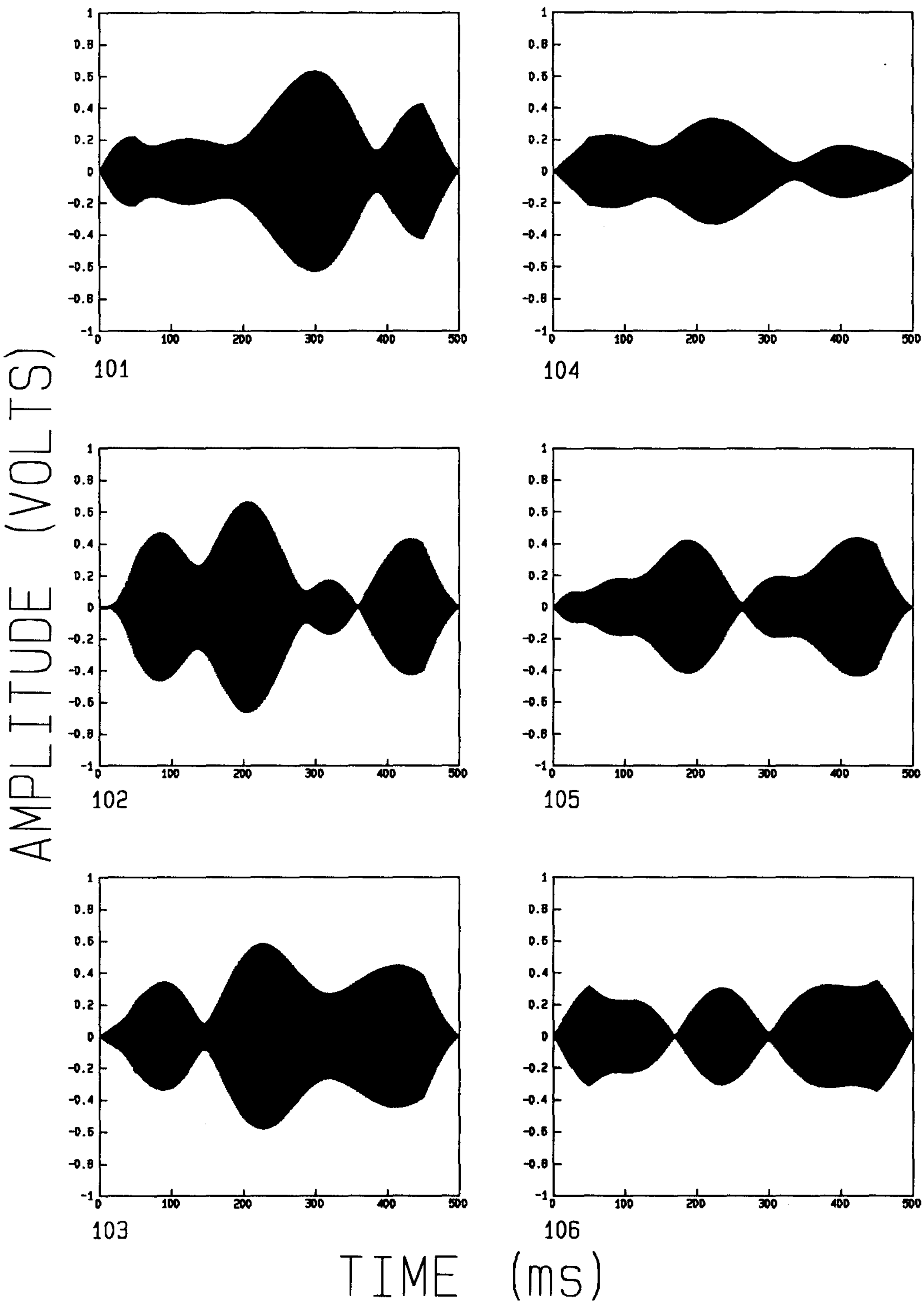

201
204
202
205
203
206
AMPLITUDE (VOLTS)
TIME (ms)

Figure 13. The time-domain waveforms of the target stimuli used in the 10-Hz bandwidth conditions are shown. Each stimulus is 500 ms in duration with 50-ms linear onsets and offsets. The amplitudes of the waveforms are scaled identically, and shown as the output voltage of the digital-to-analog converters.


AMPLITUDE (VOLTS)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
100
200
300
400
500
101
102
103
104
105
106
TIME (ms)

GENERAL DISCUSSION AND CONCLUSIONS

The main conclusions of the present series of experiments can be summarized as follows:

1) Binaural interference occurs across frequencies when listeners attempt to lateralize a target sinusoidal component in the presence of a number of distractor components, resulting in increased thershold IDT's relative to that for the target in isolation. This interference can be greatly reduced by turning off the distractors for as little as 10 ms. Threshold IDT's approach that for the target in isolation when the distractors are turned off for 20 ms or more. Thresholds are roughly equal whether the target is on for the entire listening interval or only during the temporal notch in the distractors.

2) If two pure tones are trilled, which previous research has shown leads to the perception of two streams of pulses with differenct frequencies (Bregman, 1990) and the tones have different IDT's, thresholds are comparable to those measured for the pulses of each frequency in isolation. This is also the case when the pulses temporally overlap for all but 25 ms of each pulse. Large amounts of binaural interference are observed once again when the pulses completely overlap in time.

3) At narrow frequency spacings (100 Hz or less), the ability to detect the interaural delay of a 20-Hz wide band of noise is dependent upon the relationship between its envelope and the envelopes of two

flanking 20-Hz wide distractor bands. When the envelopes are identical, detectability is lowest for a given IDT. Detectability of the target IDT is higher when the target band has a different envelope from the distractors.

4) At wider frequency spacings (200 Hz) or at narrower target and distractor bandwidths (10 Hz), detectability of a given IDT seems to be dependent on the specific envelope of the target and distractors in an additive fashion. Certain envelopes seem to dominate over other envelopes in terms of their binaural information, with the effect that if an envelope applied to the target produces high detectability of the target IDT, when the same envelope is applied to the distractors, it results in low detectability of the target IDT. Envelopes that are more "dominant" seem to have lower envelope variability, although very low envelope variability appears to diminish this dominance.

These experiments have shown that lateralization of a target sinusoidal tone or band of noise in the presence of distractor frequencies is facilitated by a brief isolated presentation of the target. The manipulations of the stimulus that resulted in an isolated presentation of the target (turning off the distractors, trilling the tones, or using different envelopes for target and distractors) have been identified in previous research as cues for auditory stream segregation or the perception of several simultaneous auditory events. However, producing the perception of a separate auditory event is not sufficient to eliminate the binaural interference produced across frequencies, as observed in the author's previous research in which an onset asynchrony between the target and distractors was introduced

(Stellmack and Dye, 1989). Although the target appeared to stand out in terms of its pitch against the background of distractors, threshold IDT's for the target among distractors were still much higher than those for the target in isolation. In this way, the results of the present series of experiments provide further support for the notion put forth by Dye (1990) and Stellmack (1990) that binaural information is combined across frequencies in certain situations.

In Experiment III, large amounts of interference occurred when the target and distractors had coherent envelopes only at frequency spacings of 100 Hz or less, or within a monaural critical band. This suggests that when binaural interference is observed at wider frequency spacings, it is to some extent "manufactured" at a higher level of processing by the auditory system when the spectral and temporal properties of the stimulus strongly suggest that the stimulus components originate from a common source. When streaming cues conflict with one another, for example, when spectral components with simultaneous onsets and offsets have different interaural information, spatial cues are usually the weakest for the segregation of auditory streams (Handel, 1989). For example, Deutsch (1975) describes an auditory illusion in which two different series of tones are presented to each ear. Rather than perceiving the tones in the correct sequence in each ear, the listener often reports that the tones are organized in each ear according to a musically logical progression of pitches that is actually occurring across the ears. Organization of the stimulus according to pitch, or melody in this case, is preferred to organization on the basis of spatial cues. It is not surprising that spatial cues are given little weight when one considers the fact that spectral incoherence can result when a single sound source emits frequencies between 500 and 2000 Hz (Kuhn, 1977). As a result, spatial cues are more unreliable as a basis of perceptual organization than pitch cues or temporal cues. When a small number of discrete spectral components are presented (from 3 to 9) with only one component interaurally delayed (as in Dye, 1990 and Stellmack, 1990), perhaps it is more reasonable for the auditory system to conclude that they originated from a common source and to combine binaural information across frequencies than to perceptually segregate the single interaurally delayed component. On the other hand, when the auditory system is confronted with a narrow band of noise consisting of many frequency components with common interaural delays and an interaurally delayed envelope, as in Experiment III of the current series of experiments, this may be sufficiently strong evidence that the band of noise was emitted from a different source than several spectrally remote distractor bands, resulting in spectrally analytic binaural processing. The binaural interference observed when the noise bands fall within a monaural critical band may reflect a physical inability of the system to separate the binaural information.

It would appear then that the combination of interaural information across frequency spacings greater than a monaural critical band results not from a lack of frequency resolution in the binaural system, but rather from a weighing of evidence at a higher level of processing that concludes that remote spectral frequencies were produced by a common source. Recent research demonstrating spectrally synthetic binaural processing of two frequency components by most listeners (Dye and Stellmack, 1992) may be another example of pitch and temporal cues dominating over spatial cues.

In dichotic pitch experiments (e.g. Cramer and Huggins, 1958, and Yost, 1991b), when a subset of components of a broadband noise are interaurally delayed, listeners report perceiving a pitch corresponding to the center frequency of the interaurally shifted band and that the pitch occupies an intracranial position separate from that of the remaining background of noise. In this type of stimulus, there are apparently no pitch or temporal cues available. The stimuli presented to the ears are simply broadband noises which, by themselves, produce no perception of pitch. However, the interaural differences that result in dichotic pitches are much larger than the interaural differences that can be detected in isolation for the same interaurally delayed band. It seems that this is an another example in which weak spatial cues must be made relatively large in order for perceptual organization on the basis of those cues to occur. In fact, dichotic pitches are usually heard most easily when the interaural difference is abruptly introduced after a presentation of at least 500 ms of diotic noise (Yost, 1985). In this case, the temporal cue is consistent with the spatial cue and serves to enhance segregation of the interaurally delayed band.

The present series of experiments has shown that spatial cues for the segregation of auditory objects are relatively weak compared to spectral and temporal cues. When temporal cues support segregation of the target component and the target component appears briefly in isolation during the course of the stimulus presentation, interaural delays of the target are nearly as detectable as when the target is presented in isolation. Temporal modulation of a stimulus of the sort that is present in narrow bands of noise influences the potency of spatial information. An as-yet-unspecified property of the temporal envelopes of auditory stimuli results in differential effects of different envelopes on the strength of spatial information. Spectrally synthetic processing of binaural information occurs for narrow bands of noise when stimuli temporally overlap within a monaural critical band. In general, when spatial cues are placed in competition with spectral and temporal cues, organization of the auditory world is least likely to occur on the basis of spatial cues. However, because these cues usually support one another rather than compete with one another within a given auditory stimulus in the real world, the relative weakness of spatial cues is usually not a problem. Only in the artificial setting of the laboratory do the shortcomings of binaural processing become evident.

REFERENCES

Blauert, J., and Lindemann, W. (1986). "Spatial mapping of intracranial auditory events for various degrees of interaural coherence," J. Acoust. Soc. Am. 79, 806-813.

Bregman, A.S. (1990). Auditory Scene Analysis. The MIT Press, Cambridge, MA.

Buell, T.N., and Hafter, E.R. (1991). "Combination of binaural information across frequency bands," J. Acoust. Soc. Am. 90, 1894-1900.

Cramer, E.M., and Huggins, W.H. (1958). "Creation of pitch through binaural interaction," J. Acoust. Soc. Am. 30, 413-417.

Deutsch, D. (1975). "Two-channel listening to musical scales," J. Acoust. Soc. Am., 57, 1156-1160.

Dye, R.H. (1990). "The combination of interaural information across frequencies: Lateralization on the basis of interaural delay," J. Acoust. Soc. Am. 88, 2159-2170.

Dye, R.H., and Stellmack, M.A. (1992). "A technique for assessing spectrally analytic and synthetic binaural processing," J. Acoust. Soc. Am. (submitted abstract).

Grantham, D.W., and Wightman, F.L. (1978). "Detectability of varying interaural temporal differences," J. Acoust. Soc. Am. 63, 511-523.

Handel, S. (1989). Listening. The MIT Press, Cambridge, MA.

Hartmann, W.M. (1987). "Temporal fluctuations and the discrimination of spectrally dense signals by human listeners," in Auditory Processing of Complex Sounds (W.A. Yost and C.S. Watson, eds.), Lawrence Earlbaum Associates, Hillsdale, N.J.

Hartmann, W.M. (1988). "Pitch perception and the segregation and integration of auditory entities," in Auditory Function: Neurobiological Bases of Hearing (G.M. Edelman, W.E. Gall, and W.M. Cowan, eds.), John Wiley and Sons, New York, N.Y.

Henning, G.B. (1980). "Some observations on the lateralization of complex waveforms," J. Acoust. Soc. Am. 68, 446-454.

Judd, T. (1979). "Comments on Deutsch's musical scale illusion," Perception & Psychophysics, 26, 85-92.

Kollmeier, B., and Gilkey, R.H. (1990). "Binaural forward and backward masking: Evidence for sluggishness in binaural detection," J. Acoust. Soc. Am. 87, 1709-1719.

Kuhn, G.F. (1977). "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am. 62, 157-167.

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am., 49, 467-477.

McFadden, D. (1966). "Masking-level differences with continuous and with burst masking noise," J. Acoust. Soc. Am. 40, 1414-1419.

McFadden, D., and Pasanen, E. (1976). "Lateralization at high frequencies based on interaural time differences," J. Acoust. Soc. Am. 59, 634-639.

Moore, B.C.J., and Glasberg, B.R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am. 74, 750-753.

Perrott, D.R. (1984). "Concurrent minimum audible angle: A re-examination of the concept of auditory spatial acuity," J. Acoust. Soc. Am. 75, 1201-1206.

Rice, S.O. (1954). "Mathematical Analysis of Random Noise," in Selected Papers on Noise and Stochastic Processes (N. Wax, ed.), Dover Publications, Inc, New York, NY.

Stellmack, M.A. (1990). "The combination of interaural information across frequencies: The effects of number and spacing of components, onset asynchrony, and harmonicity," (unpublished Master's Thesis), Loyola University of Chicago.

Stellmack, M.A., and Dye, R.H. (1989). "The effect of onset asynchronies between components on the ability to detect a delayed component embedded in a diotic complex," J. Acoust. Soc. Am. Suppl. 1, 86, S11.

Stellmack, M.A., Dye, R.H., and Jakubczak, S.V. (1989). "The effect of number of components and component spacing on the ability to detect a delayed component embedded in a diotic complex," J. Acoust. Soc. Am. Suppl. 1, 85, S83.

Trahiotis, C., and Bernstein, L.R. (1990). "Detectability of interaural delays over select spectral regions: Effects of flanking noise," J. Acoust. Soc. Am. 87, 810-813.

Woods, W.S., and Colburn, H.S. (1992). "Test of a model of auditory object formation using intensity and ITD discrimination," J. Acoust. Soc. Am., in press.

Yost, W.A. (1985). "Prior stimulation and the masking-level difference," J. Acoust. Soc. Am. 78, 901-907.

Yost, W.A. (1991a). "Auditory image perception and analysis: The basis for hearing," Hearing Research 56, 8-18.

Yost, W.A. (1991b). "Thresholds for segregating a narrow-band from a broadband noise based on interaural phase and level differences," J. Acoust. Soc. Am. 89, 838-844.

Yost, W.A., Turner, R., and Bergert, B. (1974). "Comparison among four psychophysical procedures used in lateralization," Perception & Psychophysics 15, 483-487.

Zurek, P.M. (1985). "Spectral dominance in sensitivity to interaural delay for broadband stimuli," J. Acoust. Soc. Am. Suppl. 1, 78, S18.

APPENDIX

Target Band Frequencies and Starting Phases Used in Experiment III

	Starting Phases for Envelope:					
Frequency	201	202	203	204	205	206
739.746	-148.007	136.051	137.390	73.802	10.468	-143.726
740.356	-73.834	165.517	-151.971	-37.822	99.278	158.241
740.967	-41.494	-92.989	-17.887	-109.675	-4.222	-90.956
741.577	173.327	-37.264	-16.950	-20.958	-7.571	148.494
742.188	-27.895	178.619	-57.295	146.559	-168.512	-8.648
742.798	-9.360	-21.128	-51.422	61.848	172.733	168.193
743.408	-48.330	158.845	-55.407	178.197	36.948	-123.330
744.019	-118.973	42.736	-14.173	-47.233	111.992	71.898
744.629	-18.165	-23.033	99.474	126.393	45.191	140.694
745.239	-136.583	-64.288	39.367	-156.587	176.830	-106.769
745.850	88.579	167.856	86.033	148.350	-125.063	30.444
746.460	-93.547	-83.579	80.362	-112.738	-142.164	-85.618
747.070	-27.831	58.576	59.951	-103.898	-120.803	172.733
747.681	8.484	173.396	-120.329	15.103	107.960	-76.345
748.291	-3.296	152.284	142.272	66.839	-172.508	17.779
748.901	36.101	14.360	-131.824	-162.443	-96.144	140.697
749.512	95.031	143.663	-136.442	-35.157	-113.760	-76.361
750.122	-48.880	-99.788	-113.609	-103.470	-71.276	179.260
750.732	-114.446	-160.610	140.165	81.722	-32.471	22.760
751.343	-172.082	59.057	-169.214	18.494	-130.334	-51.767
751.953	-121.112	66.229	-6.893	-43.196	-124.229	31.928
752.563	94.468	95.054	-2.633	-152.113	22.528	-155.129
753.174	-59.696	-67.799	31.958	174.081	76.690	139.040
753.784	152.631	85.980	117.064	103.324	69.587	-145.182
754.395	-156.692	-88.309	-164.346	-120.194	36.450	-97.676
755.005	86.052	113.646	-178.602	63.532	90.981	-112.750
755.615	159.467	64.606	166.378	-60.425	14.827	-87.098
756.226	-60.361	-148.426	-28.150	-69.616	-93.757	111.969
756.836	-95.215	12.246	127.065	11.799	-149.073	86.701
757.446	16.487	75.094	38.644	-1.886	50.982	23.630
758.057	25.837	-156.031	117.877	-5.213	23.329	-131.006
758.667	-125.401	-34.961	-123.509	110.513	55.905	81.751
759.277	-71.363	-25.156	83.049	-152.862	-145.398	-100.533
759.888	40.840	86.522	164.639	28.429	-45.051	-63.983

Frequency	Starting Phases for Envelope: 101	102	103	104	105	106
745.239	-151.515	-82.716	-104.575	176.289	140.171	70.343
745.850	-26.960	-103.866	147.911	-68.998	87.577	-137.132
746.460	-34.967	-106.844	-125.284	-45.253	-111.920	52.962
747.070	-24.372	138.165	82.226	92.755	-39.107	-36.158
747.681	-124.482	-114.161	-103.167	2.541	177.490	-156.012
748.291	116.840	62.829	59.726	56.035	48.692	134.059
748.901	75.328	82.554	-2.475	175.450	-165.485	-55.446
749.512	-21.986	80.644	-8.338	-38.435	-116.684	2.050
750.122	-3.527	34.217	-87.983	109.319	82.970	66.090
750.732	-66.499	-98.951	-161.646	-95.323	168.242	111.013
751.343	72.923	132.017	60.513	67.857	-45.061	40.871
751.953	-86.931	-96.397	19.375	-26.333	10.141	-178.752
752.563	88.453	58.211	-19.069	11.521	-175.954	-64.585
753.174	-28.280	-159.668	101.251	27.836	-26.030	70.109
753.784	-30.960	22.613	-89.150	-85.466	-161.311	4.801
754.395	-120.921	-2.979	-171.010	99.574	174.365	75.109
755.005	44.172	-62.213	-87.155	-165.912	-138.737	-3.332

APPROVAL SHEET

The dissertation submitted by Mark A. Stellmack has been read and approved by the following committee:

Dr. William A. Yost, Director
Director and Professor of Hearing Sciences
Loyola University of Chicago

Dr. Raymond H. Dye, Jr.
Associate Professor, Psychology
Loyola University of Chicago

Dr. Richard R. Fay
Professor, Psychology
Loyola University of Chicago

Dr. R. Scott Tindale
Assistant Professor, Psychology
Loyola University of Chicago

Dr. Stanley Sheft
Research Assistant Professor, Parmly Hearing Institute
Loyola University of Chicago

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

April 20, 1992
Date

William A. Yost
Director's Signature