The Combination of Interaural Information across Frequencies: The Effects of Number of Components, Component Spacing, and Onset Asynchrony

Mark A. Stellmack
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_theses

Part of the Psychology Commons

Recommended Citation

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1990 Mark A. Stellmack
The Combination of Interaural Information Across Frequencies:
The Effects of Number of Components,
Component Spacing, and Onset Asynchrony

by

Mark A. Stellmack

A Thesis Submitted to the Faculty of the Graduate School
of Loyola University of Chicago in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts
April
1990
Copyright 1990, by Mark A. Stellmack
I would like to thank Toby Dye for his guidance and support on this project and for keeping me from writing nonsense.

I would also like to thank Bill Yost for his helpful comments and suggestions and for making it possible for me to conduct my research at the Parmly Hearing Institute.

I would like to thank Stan Sheft for his comments and suggestions during all phases of the research and writing involved in this project.

Special thanks to Julie, whose encouragement and support helped make this possible. This is for you.
VITA

The author, Mark Andrew Stellmack, is the son of Theodore Andrew Stellmack and Jeannette Marie (Lobatch) Stellmack. He was born March 24, 1964, in Chicago, Illinois.

His elementary education was obtained at St. Francis Borgia School, Chicago, Illinois. His secondary education was completed in 1982 at Holy Cross High School, River Grove, 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, enabling him to complete the Master of Arts in 1990. He is currently a member of the Acoustical Society of America.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>EXPERIMENT I</td>
<td>8</td>
</tr>
<tr>
<td>Methods I</td>
<td>9</td>
</tr>
<tr>
<td>Results I</td>
<td>12</td>
</tr>
<tr>
<td>Discussion I</td>
<td>15</td>
</tr>
<tr>
<td>EXPERIMENT II</td>
<td>18</td>
</tr>
<tr>
<td>Methods II</td>
<td>19</td>
</tr>
<tr>
<td>Results II and Discussion II</td>
<td>20</td>
</tr>
<tr>
<td>EXPERIMENT III</td>
<td>24</td>
</tr>
<tr>
<td>Methods III</td>
<td>25</td>
</tr>
<tr>
<td>Results III and Discussion III</td>
<td>29</td>
</tr>
<tr>
<td>EXPERIMENT IV</td>
<td>34</td>
</tr>
<tr>
<td>Methods IV</td>
<td>36</td>
</tr>
<tr>
<td>Results IV and Discussion IV</td>
<td>37</td>
</tr>
<tr>
<td>GENERAL DISCUSSION AND CONCLUSIONS</td>
<td>42</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>48</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure                                                                                           Page
1. Threshold IDT as a function of component frequency spacing and number of components, 3 subjects........ 14
2. Threshold IDT as a function of component frequency spacing for 9 components in a left-right task, 3 subjects........ 23
3. Depiction of stimuli used in Experiments III and IV.... 28
4. Threshold IDT as a function of onset asynchrony and frequency spacing, 2 subjects................................. 33
5. Threshold IDT as a function of onset asynchrony for components with different harmonic relationships, 4 subjects................................................................. 41
INTRODUCTION

Sound sources in the real world produce differences in time and intensity between the waveforms received at 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 waveform reaches the ears at slightly different times. Interaural differences of intensity (IDI's) are produced when the waveform arriving at the ear furthest from the sound source is attenuated by the head. These interaural cues are used to identify the location of the sound source on the azimuthal plane.

When a pure tone is presented to a subject over loudspeakers, both IDT's and IDI'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 waveforms 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 to 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.

Under headphones, the effects of IDT's can be separated from those of IDI's by presenting stimuli in which different spectral components have different interaural delays. The focus of the present paper is on
lateralization of spectrally incoherent stimuli.

Some effects of spectral incoherence on lateralization have been observed in experiments involving pure group delays, or delays of the stimulus envelope. For complex stimuli presented through headphones, it is possible to generate pure group delays in which the envelope at one ear is delayed relative to the envelope at the other ear, while the fine structure at the two ears is synchronous (Henning, 1980). For these stimuli, the interaural phase difference is given by $2\pi\Delta t(f_c-f)$, where $\Delta t$ is the interaural delay of the envelope and $f_c$ is the center frequency of the complex. In other words, frequencies below $f_c$ are advanced and frequencies above $f_c$ are delayed in the channel in which the envelope is delayed relative to the undelayed channel.

Many authors have demonstrated the sensitivity of the binaural auditory system to group delays (envelope delays) when the spectrum of the signal consists of only high frequencies (Henning, 1974; 1980; Henning and Ashton, 1981; Nuetzel and Hafter, 1976; Bernstein and Trahiotis, 1985). Although the auditory system appears to be unable to follow the temporal fluctuations of the fine structure when the carrier is above approximately 2 kHz, it can phase-lock to the envelopes, which are compared at the two ears. Thus, lateralizations of high-frequency waveforms can occur on the basis of interaural delays of the envelope. However, some evidence suggests that sinusoidally amplitude-modulated low-frequency tones can not be lateralized on the basis of these envelope delays (Henning, 1980; Henning and Ashton, 1981). (While Bernstein and Trahiotis, 1985, demonstrated some sensitivity of the binaural system to envelope delays in low-frequency complexes, the
lateralization of these stimuli was found to be dominated by delays in the fine structure.) For envelope delays of sinusoidally amplitude-modulated low-frequency tones, the lower sideband is advanced in phase and the upper sideband is delayed. If listeners were able to perform a component-by-component analysis, the lower sideband would be lateralized to one side, the upper sideband to the other, and the diotic carrier would appear to be centered. In this case, even if envelope processing of amplitude-modulated low-frequency carriers were not possible, one might expect that Henning’s subjects could have been able to perform the lateralization task by attending to one sideband or the other, particularly since the frequency spacing between components exceeded estimated critical bandwidths and the interaural delays of the upper and lower sidebands would have been easily detectable in isolation.

Dye (1990) suggested two possibilities for explaining Henning’s results for AM waveforms with low frequency carriers. First, subjects may be unable to lateralize these waveforms due to general confusion, because different frequency components move in different directions in the two intervals of the two-alternative forced choice task employed by Henning. A second possibility raised is that the binaural auditory system combines interaural information across the frequency domain, such that a single intracranial image is formed at some "average" position based on the relative interaural delays of the individual components. Dye further examined the latter explanation.

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 a left-right judgment 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 approached 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. (A similar peak in thresholds at $\Delta f = 50$ Hz was seen in the data from the present experiments.)

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 2-IFC 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 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 large 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.

The present series of experiments addresses 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. That is, the focus of these experiments is the simple detection of the delay rather than identification of the direction of the delay. In Experiment I, the effects of the number of components and component spacing on the detection of an interaural delay of a target component are examined. In Experiment II, subjects must identify the direction, left or right, of the interaural delay of a single component in some of the multi-component complexes used in Experiment I. Experiment III examines the effects on threshold IDT's of an onset asynchrony between the delayed component and diotic components in a complex. In Experiment IV, subjects must detect the interaural delay of a single component among diotic components with various harmonic and inharmonic relationships, and with various onset asynchronies between the delayed and diotic components.
EXPERIMENT I

Threshold interaural differences of time (IDT's) were measured for multi-component complexes centered at 753 Hz as a function of number of components (3, 5, 7, 9) and component spacing (10, 25, 50, 100, 200 Hz). Subjects judged whether or not the 753-Hz component of the complex was interaurally delayed on each trial. If a small interaural delay of a single component of a complex produces a broadened intracranial image, simple detection of the delay would be expected to result in lower thresholds than the identification of the apparent direction of movement of the complex (as in Dye, 1990). In addition, as the number of components is increased and the frequency spacing is decreased, the stimulus might be expected to approach a dichotic pitch type of stimulus. A dichotic pitch stimulus is a stimulus presented over headphones that consists of either a band of noise in which the phase of a narrow range of frequencies has been shifted in one ear relative to the other (Cramer and Huggins, 1958), or, similarly, a number of discrete spectral components in which some subset (usually one) of those components contains an interaural delay (Kubovy and Daniel, 1983). A pitch is usually heard that corresponds to the center frequency of the shifted region. In the present experiment, such a cue might become available for stimuli consisting of large numbers of components and narrow frequency spacings. The effect of dichotic pitch cues would be to lower threshold IDT's.
A modified same-different task was used in which a diotic 753-Hz pure tone was presented in the first interval (the "cue tone") followed by the test stimulus in the second interval. Both intervals were 200 ms in duration with 10 ms linear rise/decay times and were separated by 300 ms. The cue tone was intended to help subjects identify the perceptual "center" and to indicate which tone would be delayed in case subjects were able to attend to that particular pitch and its apparent position.

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. Thresholds were also obtained for a 753-Hz tone in isolation, in which case only that tone was present in the test interval. This task is referred to as a "left-center" task because the interaural delay, when present, resulted in an image to the left of the midline for a pure tone, while a diotic pure tone would appear 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. 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.

Thresholds were estimated from 3- or 4-point psychometric functions, with each point based upon at least 100 trials. The threshold interaural delay was defined as the delay yielding $d' = 1.00$.

Stimuli were presented through Telephonics (TDH-39) 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. The levels of the signals were adjusted with variable attenuators (Tech Lab, Inc.) before being passed on to Crown stereo amplifiers 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 300-500 trials per subject.

The three observers in Experiment I were the author and two co-
workers, all of whom 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.
RESULTS I

Figure 1 shows the results for each of the three subjects. Threshold IDT's are plotted against frequency separation, with number of components represented by the different symbols. Each panel shows data from a single subject. The dashed horizontal line near the bottom of each panel is that subject's threshold for a 753-Hz tone in isolation.

First, 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 in all conditions run. All thresholds for the complex stimuli are higher than those for the 753-Hz tone in isolation and thresholds increase with increasing number of components. Second, the 3-component thresholds are approximately equal across Δf's for each subject. However, for 5, 7, and 9 components, subjects TD and MS have highest thresholds at Δf = 50 Hz and subject SJ has highest thresholds at Δf = 25 Hz, with decreasing thresholds at wider and narrower frequency spacings. The non-monotonic functions obtained suggest that different perceptual cues may be available at high and low frequency spacings, with a transition from the use of one cue to the other occurring at a Δf around 25-50 Hz.
Figure 1. Threshold interaural differences of time are plotted as a function of the frequency spacing between components (in Hz). All complexes consisted of equally-spaced components centered at 753 Hz. Each symbol represents a different number of components, as indicated. Each panel in the figure shows data from a single subject.
DISCUSSION I

One can imagine three ways in which the stimuli in Experiment I may be perceived:

1) Listeners "hear-out" the single delayed component, that is, the target component (753 Hz) is perceived at some intracranial position off the midline while the diotic distractors are perceived at the midline. Dye (1990) rejected this possibility upon observing that subjects were unable to identify which of two components of a complex was delayed across the two intervals of a 2AFC experiment. On the other hand, dichotic pitch stimuli consisting of large interaural delays in a subset of a broad band of spectral components often produce a pitch associated with the center frequency of the delayed band at some lateral position other than that of the remainder of the components (e.g. Yost, Harder, and Dye, 1987). In addition, Kubovy and Daniel (1983) provide a demonstration in which the 8 tones of a musical scale are presented simultaneously with one tone interaurally shifted in phase by 45°. Once again, one can hear out the pitch associated with the delayed component at a separate intracranial position from the remainder of the components. In this demonstration, the pitch of the delayed component stands out so clearly that a melody can be played by varying which component is delayed.

2) Interaural information is combined across frequencies such that the entire complex is lateralized at a single intracranial position
based on some weighted average of the interaural delays of the individual components. This "synthetic" processing of binaural information was suggested by Dye's (1990) results but is not supported by results obtained with dichotic pitch stimuli.

3) When the target component is interaurally delayed by small amounts, the intracranial image does not change position as a whole but merely broadens relative to that of the completely diotic stimulus (shown for noise stimuli by Blauert and Lindemann, 1986). As stated previously, this possibility is not ruled out by Dye's results in that the delays used in his experiment may have produced a single broad intracranial image. The direction in which the breadth of the image changed may have been detected but the components were not perceptually segregated, preventing subjects from identifying which was delayed. For dichotic pitch stimuli, the interaural delays are sufficiently large that the delayed frequencies occupy a different intracranial position from the undelayed frequencies.

Perhaps one of the three perceptual events described above occurs at wide frequency spacings and another occurs at narrow spacings resulting in the peaked functions shown in Figure 1. The three observers who participated in Experiment I agreed that it seemed as if interaural delays of the 753-Hz component at large frequency spacings resulted in a shifting of the intracranial image from the midline while at narrow spacings the delay produced a broadening of the image with little or no change in position. In order to examine this possibility, some of the conditions of Experiment I, a "left-center" task, were repeated in Experiment II as a "left-right" task. In other words, in
the test interval the target component was delayed either to the left or right. This is similar to Dye (1990) in that subjects identified the direction of the delay, rather than simply detecting the delay as in Experiment I.
EXPERIMENT II

If movement of the intracranial image occurs at wide frequency spacings, one would expect threshold IDT's obtained in this experiment at these spacings to be similar to those obtained in Experiment I. On the other hand, if the intracranial image merely broadens at narrow spacings, it would be expected that threshold IDT's at these spacings would be higher than those obtained in Experiment I. In the left-center task (Experiment I), the change in size of the intracranial image can be used by subjects to perform the task. In this experiment, the stimulus in the test interval always contains an interaurally-delayed component, so the size of the intracranial image cannot be used as a cue. This assumes that the change in size of the image at narrow frequency spacings provides some type of directional cue as the interaural delay becomes sufficiently large, or subjects could not perform this task at all.
METHODS II

Threshold IDT's were measured in a left-right task for 9-component complexes centered at 753 Hz with frequency spacings of 10, 25, 50, and 100 Hz. A 753-Hz cue tone was presented in the first interval of each trial followed by a multi-component test stimulus in the second interval. In the test interval the 753-Hz component was delayed to either the right or left ear and the remaining distractor components were diotic. The subjects' task was to indicate to which side the target component was delayed.

The remaining stimulus parameters and data collection procedures were identical to those used in the last experiment. Each threshold IDT is the overall difference between the IDT's for the "left" and "right" stimuli that would produce a $d'$ of 1.00, that is, the change in IDT from the left stimulus to the right stimulus ($\Delta$IDT). For example, if a subject showed threshold performance in the left-right task when the interaural delay to the left or right was 50 $\mu$s, this would be plotted as a 100 $\mu$s threshold IDT.

The three subjects in this experiment were the three subjects from Experiment I.
RESULTS II AND DISCUSSION II

Figure 2 shows the results for each of the three subjects, with each panel representing data from a single subject. Threshold IDT’s are plotted against frequency spacing. The open symbols connected by dotted lines represent the data collected in the left-right task (Experiment II), while the filled symbols connected by solid lines represent the comparable 9-component data for these subjects in the left-center task (Experiment I).

The results for the three subjects in this experiment show much more variability across subjects than those from Experiment I. For Subject TD, it can be seen that the left-right thresholds are always larger than the left-center thresholds, although they are very similar at the widest spacing. Unlike the left-center data, thresholds in the left-right task continue to rise as the frequency spacing decreases, consistent with the suggestion that IDT’s at narrow frequency spacings result in a broadening of the intracranial image that might be used as a cue in the left-center task but not in the left-right task.

Left-right thresholds for Subject MS are higher than the left-center thresholds at all but the widest frequency spacing, once again supporting the idea that broadening of the intracranial image is a potential cue in the left-center task. However, the left-right data show a peak at a frequency spacing of 50 Hz, similar to that found in the left-center data. This finding is inconsistent with the image-
broadening hypothesis.

Unlike the other two subjects, Subject SJ produced thresholds in the left-right task that are lower than those in the left-center task. In the left-right conditions, thresholds did not show the same steep decrease at narrow frequency spacings as in the left-center conditions.

In general, the results from the present experiment provide modest support for the suggestion that the cue for detecting IDT's at narrow frequency spacings is a broadening of the intracranial image. When the subject's task is to merely detect the presence of an interaural delay, this broadening of the image can be used as a cue. When the direction of the IDT must be identified, broadening of the image is not a useful cue, eliminating the steep reduction of thresholds at narrow spacings. The similarity of thresholds at wide frequency spacings in the left-right and left-center tasks suggests that these tasks are performed in the same manner at these wide spacings, perhaps with judgments based on the apparent movement of the intracranial image across intervals.
Figure 2. Threshold interaural differences of time (in µs) are plotted as a function of the frequency spacing between components (in Hz) for three subjects. All complexes consisted of 9 components centered at 753 Hz. In the left-right conditions (open symbols), subjects discriminated between a complex in which the center component was delayed to the left and a complex in which the same component was delayed to the right. In the left-center conditions (filled symbols), subjects discriminated between a complex in which the center component was delayed to the right and a completely diotic complex.
EXPERIMENT III

In experiments in which spectrally analytic binaural processing occurs, it is often observed that dynamic changes in the interaural configuration of the stimulus is an important factor. In the demonstration by Kubovy and Daniel (1983) discussed previously, in which one can clearly hear out an interaurally delayed component in an 8-component complex, the interaural delay is applied to different components over the duration of the stimulus. (This is also the case in Yost et al., 1987.) The following experiment examines the ability of listeners to lateralize an interaurally delayed target component in a multi-component complex when an onset asynchrony exists between the target and the diotic distractors.

In Experiments I and II, interaural delays of the target component were achieved by shifting the phase of the target in one channel relative to the other. In this experiment, interaural delays were generated in the same manner for some conditions, but several additional conditions were run in which the interaural delay of the target component was "real" in the sense that the target was gated on in one channel before the other, resulting in both an onset and ongoing interaural delay.
METHODS III

The method and equipment were similar to Experiment I. Once again, a diotic 200 ms, 753-Hz cue tone was presented in the first interval of each trial followed by a multi-component test stimulus in the second interval. Unlike Experiment I, onset asynchronies between the target and distractor components of 0, 50, 100, or 200 ms were generated, with the diotic distractor components gated on first, as shown in Figure 3. The duration of the target was always 200 ms, and the duration of the distractors was such that they remained on for the duration of the target (all components were gated off simultaneously). Thus, the total duration of the test interval was 200 ms plus the onset asynchrony. (An onset asynchrony of 0 ms duplicates the conditions of Experiment I.)

Threshold IDT's were measured for 7-component complexes centered at 753 Hz with frequency spacings of 10, 50, and 100 Hz. The interaural delay of the target component at these spacings was achieved by advancing the target's phase in one channel relative to the other.

Additional thresholds were obtained for $\Delta f = 100$ Hz in which the delay of the target was real, that is, the target was gated on in one channel before being gated on in the other. (In these cases, the onset asynchrony between components is the time between the gating-on of the distractors and the gating-on of the channel containing the advanced target.) In these conditions, however, all components in both channels
were gated off simultaneously. Thresholds were computed for onset
asynchronies between the target and distractors of 0, 25, 50, 100, and
200 ms for the conditions in which the target component had a real
interaural delay. The 25 ms onset asynchrony was added to this stimulus
condition because it is slightly larger than the onset asynchrony at
which listeners can discriminate which of two different stimuli preceded
the other (Hirsh, 1959). For the conditions in which the target
component contained a real interaural delay, the output rate of the
digital-to-analog converters was set to 100 kHz to allow for interaural
delays that are multiples of 10 µs.

The test interval was either a completely diotic complex or the
target was interaurally delayed to the right (both cases included the
appropriate onset asynchrony between components). Thus, the subjects’
task was to indicate whether or not the target component contained an
interaural delay in the test interval (a left-center task).

Component rise/decay times were 10 ms, and all components were
equal in amplitude (55 dB SPL). The remaining stimulus parameters and
data collection procedures were identical to those used in Experiment
I.

Subjects MS and SJ in this experiment were also participants in
Experiments I and II.
Figure 3. This figure depicts the stimuli presented in each trial of Experiments III and IV. The cross-hatched areas represent intervals during which the target component was present. The cue tone consisted of the diotic target component alone. In the test interval, the distractor components were gated on first, followed by the target component after a particular onset asynchrony (O.A.). In the test interval, the target component was either diotic or interaurally delayed to the right (a left-center task). The distractors remained on for the entire test interval and all components were gated off simultaneously.
RESULTS III AND DISCUSSION III

The results for each of the subjects in this experiment are shown in Figure 4. Threshold IDT's are plotted against onset asynchrony, with component frequency separation represented by the different symbols. (All test stimuli were 7-component complexes.) The open symbols represent those conditions in which interaural delays of the target component were achieved by shifting its phase in one channel, and the filled symbols are data from conditions in which the target had a real delay.

It is clear that the onset asynchrony between the target and distractor components had no effect on the ability of Subject MS to lateralize the target components. The functions are nearly flat across the different onset asynchrony conditions. The results for this subject are consistent with those of Experiment I in that highest thresholds are obtained at $\Delta f = 50$ Hz. For this subject, thresholds are slightly lower for the real delay conditions relative to the phase delay conditions for $\Delta f = 100$ Hz, but onset asynchronies of the target still did not lead to improved performance relative to the 0 ms onset asynchrony in the real delay conditions.

For Subject SJ, the onset asynchronies had essentially no effect on performance in the $\Delta f = 100$ Hz conditions, and the real delay of the target did not result in any consistent improvement over the phase delay conditions across onset asynchronies. In the 10- and 50-Hz frequency
separation conditions, this subject's performance was poorer (thresholds are higher) when an onset asynchrony between the target and distractors is introduced relative to the no-onset-asynchrony conditions. For $\Delta f = 10$ Hz, thresholds increase with increasing onset asynchrony, so that the threshold at 200 ms onset asynchrony is over 50% larger than the threshold obtained with no onset asynchrony. The function for $\Delta f = 50$ Hz is somewhat more complicated, with the thresholds increasing at 50 ms onset asynchrony, decreasing at 100 ms onset asynchrony, and increasing again at 200 ms onset asynchrony. It appears that at the smaller $\Delta f$'s and when the diotic distractors were on for longer durations before the target, the distractors had a stronger influence on this subject's lateralization of the complex in the form of increased interference.

Although the data for the two subjects are dissimilar, it can be concluded that onset asynchronies between the target and distractors did not lead to reduced thresholds relative to conditions in which all components were gated on simultaneously, that is, synthetic binaural processing of these stimuli appears to occur. In addition, a real delay of the target component results in little or no improvement over the phase delay conditions. It should be noted that the subjects agreed that the onset asynchrony between components made the pitch of the target component appear to stand out from the distractors, but this effect did not lead to a spatially distinct lateralization of the target.

It is possible that the onset asynchronies used in this experiment were not large enough to eliminate the interference produced by the
diotic distractor components. In a study of masking level differences (MLD's), Yost (1985) found that the diotic noise background had to precede the onset of the interaurally phase shifted tone by at least 500 ms in order to obtain the maximum MLD. MLD's in that experiment gradually decreased as the onset asynchrony between the noise background and interaurally delayed tone was made smaller. Although the maximum onset asynchrony in the present experiment was only 200 ms, Subject MS showed no effect of changing onset asynchrony, and thresholds for Subject SJ actually decreased with decreasing onset asynchrony in the 10 Hz and 50 Hz frequency spacing conditions, in contrast with Yost's (1985) data. Therefore, it is not clear what the effect of larger onset asynchronies may have been in the present experiment.
Figure 4. Threshold interaural differences of time (in µs) are plotted as a function of the onset asynchrony (in ms) between the distractor components and target component. All complexes consisted of 7 equally-spaced components centered at 753 Hz. Each symbol represents a different frequency spacing in Hz. The open symbols show data from conditions in which the interaural delay of the target component was achieved by shifting its phase in one channel relative to the other. The filled symbols (real delay) show data from conditions in which the interaural delay of the target component was achieved by gating the target on in one channel before the other, resulting in both an interaural onset delay and ongoing phase disparity.
Subject MS

- 10 Hz
- 50 Hz
- 100 Hz
- 100 Hz, real delay

Subject SJ
EXPERIMENT IV

If binaural processing of the stimuli in the previous experiments had occurred in a spectrally analytic manner, subjects would have heard the diotic distractor components as a group at an intracranial position near the midsaggital plane and the target component at some lateral position consistent with its interaural delay. It might be expected that if the target component could be made more perceptually distinct relative to the distractor components, spectrally analytic binaural processing may be seen. Although the onset asynchronies between components in Experiment III made the pitch of the target more prominent, binaural processing of the inharmonic complexes was spectrally synthetic. In the following experiment, different harmonic relationships between the target component and distractors are employed (in addition to various onset asynchronies between components) as an attempt to allow spectrally analytic binaural processing of these stimuli to occur.

It has been demonstrated that the auditory system tends to fuse harmonically related components into a single entity, while a mistuned harmonic will stand out against a background of harmonic components (Rasch, 1978; Martens, 1984). Perhaps this effect can be exploited by the auditory system in order to perform spectrally analytic binaural processing of the stimulus. One set of conditions in the following experiment requires listeners to lateralize a 691-Hz component that is
presented against a diotic background consisting of 600- and 800-Hz components.

In a second set of conditions, a complex consisting of 600-, 700-, and 800-Hz components, which have a common fundamental of 100 Hz, are presented with the 700-Hz component as the interaurally delayed target. One might argue that the auditory system interprets these components as arising from a single source because of their harmonic relationship, and this in turn may lead to synthetic binaural processing of the complex. However, if the diotic 600- and 800-Hz components are gated on simultaneously, their common 200-Hz fundamental might cause them to be grouped separately from an interaurally delayed 700-Hz tone that is gated on later. In the case in which such an onset asynchrony between components exists, the target might be lateralized at an intracranial position separate from the distractors, that is, binaural processing of the stimulus may proceed analytically.

To examine the possibility that a more spectrally rich background might lead to perceptual segregation of the target, in a third set of conditions, a 753-Hz target is presented among harmonically-related distractors of 500, 600, 700, 800, 900, and 1000 Hz (producing a 7-component complex). Finally, for comparison purposes, a completely inharmonic complex consisting of a 753-Hz target among 653- and 853-Hz distractors is presented.

Onset asynchronies between the target and distractors of 0, 25, and 200 ms were run for each set of conditions. Based on the results of Experiment III, only the shortest and longest non-zero onset asynchronies were deemed necessary.
METHODS IV

The method, equipment, and computation of thresholds were identical to those in Experiment III, and the stimuli were as described above. The frequency of the cue tone was equal to that of the target in each block of trials. Interaural delays of the target were achieved by advancing the phase of the target in the left channel. The target was either interaurally delayed or diotic in each test interval, and the subjects' task was to detect the delay (a left-center task).

Subjects MS and SJ in this experiment were also participants in Experiments I, II, and III. Subjects NS and AS were undergraduate students at the author's university and were paid to participate in the experiment. Subjects NS and AS practiced lateralizing various stimuli in approximately 2000 trials over the course of two weeks before data were collected for this experiment.
RESULTS IV AND DISCUSSION IV

The results for this experiment are shown in Figure 5. Each subject's data are plotted in a different frame of the figure. Threshold IDT's are plotted against onset asynchrony. The open symbols represent the various 3-component conditions and the stars represent the 7-component condition, as indicated in the figure. Note that the ordinate of the plots for Subjects NS and AS have a larger range than those for the other two subjects.

The results for Subject MS show that, as in Experiment III, an onset asynchrony between the target and distractors has no effect on threshold IDT's for this subject, regardless of the harmonic relationship of the components: thresholds are nearly equal at each onset asynchrony for a given complex. Thresholds for the 7-component complex are larger than those for the 3-component complexes, consistent with the observation from Experiment I that interference effects increase with increasing numbers of diotic distractors. At an onset asynchrony of 25 ms, slightly lower thresholds were obtained for the 600-691-800-Hz complex relative to the other 3-component complexes, but at a 200-ms onset asynchrony, thresholds for the 3-component complexes are nearly equal. The harmonic relationship between the target and diotic distractors appears to have a minimal effect, if any, on this subject's ability to lateralize the target.

Although Subject SJ's data show greater variance between
conditions, once again an onset asynchrony between components results in little or no effect on thresholds for the 3-component conditions. Thresholds are slightly higher in the onset asynchrony conditions than in the simultaneous onset conditions for the 3-component complexes (except for the 25-ms onset asynchrony of the 600-700-800-Hz complex). Thresholds for the 7-component complex are similar to those obtained for this subject for the 7-component complexes of Experiment III: thresholds are nearly 100% higher in the onset asynchrony conditions than in the simultaneous onset condition. Thresholds for the 3-component inharmonic complex (653-753-853 Hz) are slightly higher than those for the other 3-component conditions, but it is clear that the onset asynchrony between components did not result in lower thresholds within any one condition, with the exception noted previously.

The data for Subject NS and AS are similar to those of Subject SJ. For the 3-component stimuli, thresholds are slightly higher in the onset asynchrony conditions relative to the simultaneous onset conditions, but there is little difference between the thresholds for the three different 3-component complexes at each level of onset asynchrony. Thresholds increase dramatically for these subjects with the addition of an onset asynchrony in the 7-component complex, and thresholds for the 7-component complex overall are higher than those for the 3-component complexes (as for subjects in Experiment I).

The results of this experiment are consistent with those of Experiment III in that the addition of an onset asynchrony between the target and distractors does not decrease threshold IDT's. On the contrary, thresholds often increase in the onset asynchrony conditions,
particularly with larger numbers of spectral components. It seems safe to conclude that spectrally synthetic binaural processing is not diminished by the type of onset asynchrony utilized in this experiment.

Subjects did produce slightly lower thresholds when the target was inharmonically related to the harmonic distractors (600-691-800-Hz condition), but thresholds remained many times higher than those obtained for the target in isolation (around 10 \(\mu s\) for these subjects). In other words, the diotic distractors produce considerable interference regardless of their harmonic relationship to the target.
Figure 5. Threshold interaural differences of time (in µs) are plotted as a function of the onset asynchrony (in ms) between the distractor components and target component for 4 subjects. The open symbols show data for 3-component complexes (consisting of the frequencies shown in the figure) in which the center component was the target. The stars (labelled "7 Component") show data for a complex consisting of a 753-Hz target and distractors of 500, 600, 700, 800, 900, and 1000 Hz.
GENERAL DISCUSSION AND CONCLUSIONS

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

1) The presence of 2 to 8 diotic components spaced equally 10 to 200 Hz apart and centered at 753 Hz interferes with the ability of subjects to detect an interaural difference of time in a target component at 753 Hz. This is shown by increased threshold IDT's for the interaurally delayed 753-Hz tone when other diotic components are present relative to the threshold IDT for the pure tone in isolation. The greatest interference occurs at a frequency spacing of approximately 25 to 50 Hz.

2) The amount of interference increases with increasing numbers of diotic distractor components.

3) An onset asynchrony between the diotic distractor components and the interaurally delayed target component of up to 200 ms fails to reduce the interference produced by the distractor components, and in many cases leads to even poorer performance.

4) When the interaurally delayed target component is inharmonically related to the diotic distractor components, considerable interference in the lateralization of the complex is observed, even when an onset asynchrony between the target and distractor components exists. It was observed that an onset asynchrony between the target and distractor components made the pitch of the target component appear to
stand out from the distractor components, although detection of interaural delays was not facilitated.

All of the above results are consistent with Dye's (1990) conclusion that binaural processing of multicomponent signals is spectrally synthetic for conditions in which the complexes consist of a small number of low-frequency components. Spectrally incoherent stimuli consisting of a small number of components produce a single intracranial image. In effect, the binaural system is assuming that components having different interaural information are arising from the same source. Perrott (1984) observed similar interference in the free field. Two tones differing in frequency were presented to subjects simultaneously from two different loudspeakers. It was found that the obtained concurrent minimum audible angles were several times greater than those obtained when the signals emitted from the two speakers did not temporally overlap. This effect occurred even when Δf was as large as 101 Hz.

As observed indirectly in the present series of experiments, the pitch of a particular spectral component can be made to stand out from that of a complex without introducing an interaural difference of time in the single component, in other words, by way of certain monaural cues. Subjects reported that the pitch of the target component was more prominent relative to the pitches of the distractor components when an onset asynchrony between components was introduced (also when the target component is turned off briefly; Woods, Sorkin, and Boggs, 1979; Kubovy and Daniel, 1983) or when the target component was inharmonically related to the distractor components (see also Rasch, 1978; Martens,
Thus the auditory system is spectrally analytic with respect to the pitch of a single component in a complex when the onset of that component differs from that of the remainder of the complex, or when the component is a mistuned harmonic among many harmonically related components. However, the auditory system is spectrally synthetic with respect to the binaural information in spectrally incoherent stimuli when small interaural differences of time are present. To some extent, binaural information is combined across frequencies even when the pitch of the target component is made more salient by an onset asynchrony or detuning. The auditory system is spectrally analytic for certain monaural cues such as temporal information and harmonic groupings and spectrally synthetic for binaural information at the same time.

It was pointed out that dichotic pitch cues that could have been used by the subjects in the present experiments were not produced by stimuli that might have been expected to produce such effects. This raises the question of the relationship between the results of dichotic pitch experiments and the lateralization of spectrally incoherent complexes. Dichotic pitch experiments (e.g. Cramer and Huggins, 1958, and Yost, 1990) show that the detectability of a subset of spectral components that is presented with other diotic distractor components is greatly improved when the interaural configuration of the subset differs from that of the distractor frequencies. The pitch of the interaurally phase-shifted region is heard out from the diotic noise background, and subjects frequently report hearing the pitch at some lateral position off the midline while the diotic components are heard at the midline. However, in dichotic pitch experiments, the phase-shifted region usually
contains an interaural delay that is much larger than the threshold IDT's observed in the present experiments. The original Crèmer and Huggins (1958) experiments shifted the target frequencies through 360°, with the perceived pitch corresponding to a frequency near 180°. In Yost et al. (1987), in which a region of broadband noise is interaurally delayed, subjects report hearing the dichotic pitch in more than 50% of the presentations only when the interaural delay exceeds 60° in phase (for the narrowest shifted regions, which are most comparable to the present experiments). In contrast, the largest threshold IDT obtained in Experiment I (see Figure 1) was equivalent to an interaural phase difference of approximately 54° for the target component. All other thresholds were at interaural phase differences less than 45° for the target component, and the majority were below 30° interaural phase difference. In the present series of experiments, subjects never reported that they identified the signal interval by hearing the pitch of the cued tone stand out from the rest of the complex. Subjects always reported basing their judgments on the apparent intracranial position of the stimulus. Dye (1986) presented a 3-component complex to subjects in a two-interval task in which a different component was interaurally delayed in each interval. Subjects found it extremely difficult, if not impossible, to identify which component was interaurally delayed in each interval. The lateral position of the image was apparently a more salient cue than any dichotic pitch that may have been heard.

It would appear that subjects are able to lateralize spectrally incoherent complexes at interaural delays that are smaller than those
at which dichotic pitches are heard. Yost (1990) asked subjects to perform a task similar to that in Experiment III: detect a phase-shifted portion of a broadband noise following 200 ms of diotic noise. It was assumed that dichotic pitch thresholds were being obtained. In light of the observations of the present series of experiments and those made above, it seems possible that Yost's subjects (Yost, 1990) used lateral position as a cue instead of the pitch of the shifted spectral region.

Another possible explanation of Yost's (1990) results is that significant differences in performance exist between multi-component, broadband stimuli and narrow-band stimuli consisting of small numbers of components. Trahiotis and Bernstein (1990) reported that there is virtually no increase in threshold IDT's for bands of noise when a continuous diotic background of noise is present. Their results conflict with those in Experiments III and IV, in which an onset asynchrony between the diotic distractor components and interaurally delayed target component failed to reduce the interference produced by the distractor components.

The results of the present series of experiments when considered with the results of dichotic pitch experiments suggest that when an interaural difference of phase is introduced into a subset of spectral components in a complex stimulus, there exists a continuum of effects with respect to lateralization of the complex and the perception of dichotic pitches. It appears that as the interaural difference of phase in the subset of components increases from 0°, the intracranial position of the complex is affected first. The intracranial position of the
complex as a whole appears to shift toward the ear in which the subset of components leads in phase. The apparent intracranial position of the image is based upon some weighted average of the interaural delays of the individual spectral components. As the interaural difference in phase of the subset of components is increased further, dichotic pitch effects are heard, in which case the pitch of the phase-shifted region predominates over the pitch of the remainder of the complex. When the interaural difference in phase of the subset of components becomes sufficiently large (above approximately 45°, according to Yost et al., 1987), the pitch of the phase-shifted region can be detected, but its intracranial position becomes ambiguous. As a result, the findings from the present series of experiments support the notion that the perception of dichotic pitch depends upon the lateralization of the complex stimulus, in that lateralization of the phase-shifted spectral region accompanies and precedes (occurs at smaller interaural delays than) the perception of dichotic pitches. This suggests that special care must be taken in experiments in which detection thresholds for dichotic pitches are measured to ensure that the lateral position of the resultant intracranial image is not used as a cue.
REFERENCES


APPROVAL SHEET

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

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

Dr. William A. Yost
Professor, Psychology and Otolaryngology
Loyola University of Chicago

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

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

April 10, 1990  Raymond H. Dye, Jr.
Date  Director's Signature