An Investigation of the Contribution of Auditory-Grouping to the Comodulation Masking Release Effect.

Patrick Wynn Rappold
Louisiana State University and Agricultural & Mechanical College
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
An investigation of the contribution of auditory-grouping to the comodulation masking release effect

Rappold, Patrick Wynn, Ph.D.
The Louisiana State University and Agricultural and Mechanical Col., 1991
AN INVESTIGATION OF THE CONTRIBUTION OF AUDITORY-GROUPING TO THE COMODULATION MASKING RELEASE EFFECT

A Dissertation
Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in The Department of Communication Disorders

by
Patrick W. Rappold
Louisiana State University, 1983
Louisiana State University Medical Center, 1985
August 1991
DEDICATION

To my Father, Charles B. Rappold, Jr.,

for being the person in my life I have always

been able to count on from the time I was born.

I thank him for his guidance and, especially, for

his example. Throughout my life he has encouraged

me to "think," and for that I thank him, too.

Thanks Dad, I love you.
ACKNOWLEDGEMENTS

I thank John K. Cullen, Jr., mentor and friend, for pushing me farther than I thought I could go. I thank him for his love of science, his attention to detail, and his belief that aspiring scientists should begin with fundamentals and a wide knowledge base, before focusing interest on a narrow area of wonder. I thank him also for taking much, but not all, of the magic out of audition. When all of the magic is gone, so, too, is the need for science.

I thank M. Jane Collins for her patience and for her confidence in me. I thank her for "listening" -- in every sense of the word. I thank her for helping me become a better writer and a better person. I also thank her for black beans and rice and chocolate-caramel-fudge ice cream. Thank you, Jane.

I thank Jack S. Damico for coffee and tete-a-tete during the very early morning hours at Louie’s Cafe’. What he has accomplished in his young career as an educator and scientist has been an inspiration to me, and for that I thank him, also. I thank both Jack and his wife, Sandy, for always making me feel welcome and comfortable in their home.

I thank Lee Mendoza for debate. We have argued about everything from what’s good rock-and-roll to the influence and limitations of top-down processing in audition. I thank Lee and Vijay Ratnam for the software they have written for the hearing science lab at LSU. This dissertation would have taken longer to complete if not for their programming assistance.
I thank Wendy Jumonville for serving as a subject (without pay) in all three experiments of this dissertation. She has been a real trooper, and both her and her ears have been good to me.

Although she was unable to go the distance, I thank Joelle for her early support and her willingness to spend part of her adult life with a graduate student.

Finally, I thank my family; my parents Charles and Rosalie, and my brothers Kyle and David. Their belief in me has made all the difference.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>x</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CMR Hypotheses Based on Across-Channel Comparisons</td>
<td>2</td>
</tr>
<tr>
<td>CMR Hypotheses Not Based On Across-Channel Comparisons</td>
<td>5</td>
</tr>
<tr>
<td>Auditory-Grouping Hypothesis of CMR</td>
<td>6</td>
</tr>
<tr>
<td>EXPERTIMENT 1</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>11</td>
</tr>
<tr>
<td>METHODS</td>
<td>13</td>
</tr>
<tr>
<td>Subjects</td>
<td>13</td>
</tr>
<tr>
<td>Stimuli</td>
<td>14</td>
</tr>
<tr>
<td>Construction of AM Noise-Bands</td>
<td>15</td>
</tr>
<tr>
<td>Calibration</td>
<td>19</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>19</td>
</tr>
<tr>
<td>Experimental Procedure</td>
<td>22</td>
</tr>
<tr>
<td>RESULTS</td>
<td>23</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>26</td>
</tr>
<tr>
<td>Temporal Envelope Correlation of &quot;Uncorrelated&quot; Noise-Bands and Its Possible Relationship to CMR</td>
<td>29</td>
</tr>
<tr>
<td>Signal Detection Thresholds in Correlated Noise-Bands, and CMR</td>
<td>33</td>
</tr>
<tr>
<td>FOOTNOTES FOR EXPERIMENT 1</td>
<td>35</td>
</tr>
<tr>
<td>EXPERIMENT 2</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>37</td>
</tr>
<tr>
<td>METHODS</td>
<td>37</td>
</tr>
<tr>
<td>Subjects</td>
<td>37</td>
</tr>
<tr>
<td>Stimuli and Calibration</td>
<td>38</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>38</td>
</tr>
<tr>
<td>EXPERIMENTAL PROCEDURE</td>
<td>41</td>
</tr>
<tr>
<td>Background and Rationale</td>
<td>41</td>
</tr>
<tr>
<td>Procedure</td>
<td>44</td>
</tr>
<tr>
<td>RESULTS</td>
<td>49</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>49</td>
</tr>
<tr>
<td>ISIs and Correlation Coefficients</td>
<td>52</td>
</tr>
<tr>
<td>ISIs and CMR</td>
<td>53</td>
</tr>
<tr>
<td>FOOTNOTE FOR EXPERIMENT 2</td>
<td>58</td>
</tr>
<tr>
<td>EXPERIMENT 3</td>
<td>60</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>METHODS</td>
<td>61</td>
</tr>
<tr>
<td>Stimuli and Calibration</td>
<td>61</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>61</td>
</tr>
<tr>
<td>Part A</td>
<td>61</td>
</tr>
<tr>
<td>Part B</td>
<td>64</td>
</tr>
<tr>
<td>EXPERIMENTAL PROCEDURE</td>
<td>64</td>
</tr>
<tr>
<td>Part A; Weakly Fused Condition</td>
<td>66</td>
</tr>
<tr>
<td>Part B; Strongly Fused Condition</td>
<td>66</td>
</tr>
<tr>
<td>RESULTS</td>
<td>70</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>70</td>
</tr>
<tr>
<td>FINAL DISCUSSION</td>
<td>76</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>82</td>
</tr>
<tr>
<td>APPENDICIES</td>
<td>85</td>
</tr>
<tr>
<td>VITAE</td>
<td>91</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Temporal Aspects of Noise-Bands and Pure-tone Signal</td>
<td>16</td>
</tr>
<tr>
<td>2.</td>
<td>Amplitude-Modulated Noise-Bands</td>
<td>18</td>
</tr>
<tr>
<td>3.</td>
<td>Instrumentation Diagram for Experiment 1</td>
<td>21</td>
</tr>
<tr>
<td>4.</td>
<td>Sample Tracing of Subject MC Tracking the Pure-Tone Signal</td>
<td>24</td>
</tr>
<tr>
<td>5.</td>
<td>CMR Effect</td>
<td>25</td>
</tr>
<tr>
<td>6.</td>
<td>CMR Effect For Specific Noise-Band Comparisons</td>
<td>27</td>
</tr>
<tr>
<td>7.</td>
<td>Correlation Coefficients of Uncorrelated Noise-Band Pairs</td>
<td>30</td>
</tr>
<tr>
<td>8.</td>
<td>Correlation Coefficients Vs. CMR</td>
<td>31</td>
</tr>
<tr>
<td>9.</td>
<td>Thresholds in Correlated Noise-Bands Vs. CMR</td>
<td>34</td>
</tr>
<tr>
<td>10.</td>
<td>Instrumentation Diagram for Experiment 2</td>
<td>40</td>
</tr>
<tr>
<td>11.</td>
<td>Presentation Sequence of Stimuli for Experiment 2</td>
<td>46</td>
</tr>
<tr>
<td>12.</td>
<td>Vertical-Fusion</td>
<td>47</td>
</tr>
<tr>
<td>13.</td>
<td>Horizontal-Streaming</td>
<td>48</td>
</tr>
<tr>
<td>15.</td>
<td>Horizontal-Streaming of Each Noise-Band Pair</td>
<td>51</td>
</tr>
<tr>
<td>16.</td>
<td>Correlation Coefficients VS. ISI</td>
<td>55</td>
</tr>
<tr>
<td>17.</td>
<td>CMR VS. ISI</td>
<td>56</td>
</tr>
<tr>
<td>18.</td>
<td>Instrumentation Diagram for Experiment 3 (Part A)</td>
<td>62</td>
</tr>
</tbody>
</table>
19. Instrumentation Diagram for Experiment 3 (Part B) ......................... 65
20. Presentation Sequence for Part A of Experiment 3 .......................... 68
22. Presentation Sequence for Part B of Experiment 3 .......................... 69
23. Thresholds in Weakly Fused and Strongly Fused Conditions .......... 71
24. Thresholds in Each Correlated Noise-Band Pair ............................ 72
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Audiometric data on subjects</td>
<td>86</td>
</tr>
<tr>
<td>B. ANOVA summary table for experiment 1</td>
<td>87</td>
</tr>
<tr>
<td>C. Duncan’s Multiple Range Test for experiment 1</td>
<td>88</td>
</tr>
<tr>
<td>D. ANOVA summary table for experiment 2</td>
<td>89</td>
</tr>
<tr>
<td>E. ANOVA summary table for experiment 3</td>
<td>90</td>
</tr>
</tbody>
</table>
The comodulation masking release (CMR) effect involves improved detection of a pure-tone signal in an amplitude-modulated (AM) masker by an addition of AM noise spectrally distant from the signal (i.e., outside the critical-band of the signal). The underpinning of CMR has been elusive, and despite many previous investigations of this phenomenon it remains unexplained.

A plausible explanation of CMR involves perceptual, auditory-grouping of acoustic signals. The purpose of this project was to investigate an hypothesis based on auditory-grouping to explain the CMR effect. A series of three experiments was completed towards this purpose. In the first, signal detection thresholds were obtained for a 2000-Hz pure-tone centered in narrow-band, AM maskers. Narrow-band, AM "flankers" centered at 1700-Hz were presented concurrently with the masker-band. The AM of the flanker-band was either identical to (correlated condition), or independent of (uncorrelated condition) the AM of the masker-band. Thresholds for the tone were about 3 dB better (lower) in the correlated condition, overall -- a CMR effect of about 3 dB. In comparisons of specific correlated/uncorrelated noise-band pairs the magnitude of the CMR effect was found to vary substantially. CMR magnitude appeared to be related to the degree of envelope correlation of the "uncorrelated" noise-band pairs. In experiment two, strength of vertical-fusion of correlated and uncorrelated noise-band pairs was inferred from the interstimulus-interval (ISI) necessary to "capture"
the flanker-band into a horizontal-stream. Shorter ISIs were needed to capture the
flanker-band in correlated noise-band pairs than in uncorrelated noise-band pairs.
This suggested that the strength of vertical-fusion of noise-band pairs was greater
in correlated conditions than in uncorrelated conditions. In addition, the ISIs
needed to capture the flanker-band were shorter for the "uncorrelated" noise-band
pairs where the absolute value of the correlation coefficient was largest. In the
third experiment, thresholds for the pure-tone were obtained in correlated noise-
bands. Two conditions were created, and denoted "weakly-fused" condition and
"strongly-fused" condition. According to the auditory-grouping hypothesis of
CMR, thresholds should have been better in the strongly-fused condition than in
the weakly-fused condition. A significant difference for threshold was not found,
however, between the two conditions. This finding suggests that auditory-grouping
does not play a dominant role in CMR. Alternative explanations of why a difference
for threshold was not found between the two conditions are offered, however.
INTRODUCTION

Until recently, most hearing scientists believed that detection of a monaural signal in noise could be affected by noise coded in the same auditory channel as the signal, but not by noise coded in ipsilateral channels spectrally removed from the signal channel. This belief has recently been challenged, however, by a phenomenon known as "comodulation masking release," wherein the detection threshold for a pure-tone signal centered in an amplitude modulated (AM) narrow-band masker is 3 to 8 dB better if a second-band of noise is presented simultaneously with the masker-band -- provided the temporal envelope of the second-band is identical to the envelope of the masker-band (e.g., Hall et al., 1984; McFadden, 1986; Schooneveldt and Moore, 1987; for a review see Hall, 1987). If the temporal envelope of the second-band fluctuates independently of the temporal envelope of the masker-band, there is no improvement in the detection threshold. Thus, the effect is dependent upon coherent temporal envelopes between the masker-band and the second or flanker-band. The effect was first observed by Hall, Haggard and Fernandez in 1984, at which time they coined the term "comodulation masking release (CMR)" to denote the phenomenon.

Since discovery of the CMR effect, a number of hypotheses have been offered in attempts to identify and describe its underpinning. Most of the hypotheses proposed to date fall into one of two groups. The first group
includes those hypotheses that suggest that CMR results from an across-channel comparison of energy between the auditory channel maximally stimulated by the masker-band and pure-tone signal, and the auditory channel maximally stimulated by the flanker-band. In contrast, the hypotheses of the second group are not based on across-channel comparisons. An hypothesis that falls in neither the first or second group, and has gained attention only recently, is one based on the perceptual organization of sound. This hypothesis is the main focus of this work.

**GROUP 1 HYPOTHESES**

**(ACROSS-CHANNEL COMPARISONS)**

The hypotheses of the first group differ from one another primarily by the across-channel cue being used by the auditory system to aid in detecting the presence of a tone in one of the channels. For example, one of these hypotheses suggests that the cue for tone presence is actually a detection of an across-channel difference in overall level (Hall, 1986). The addition of a tone to one of the noise-bands raises the energy level in that auditory channel. By comparing the outputs of two channels that have identical temporal envelopes, the auditory system may be capable of recognizing smaller level differences than if the two channels are not stimulated by temporal envelopes fluctuating in synchrony. This detection of a difference in level may provide the cue for tone
presence. Hall and Grose (1988) and McFadden (1986), however, measured a substantial CMR effect despite random variations in the level of the flanker-band from one stimulus interval to the next. The random variations in the level of the flanker-band should preclude use of a cue based on an across-channel difference in overall level. Therefore, these findings discredit an hypothesis based on a level cue. A second hypothesis of this group is similar to the equalization-cancellation (EC) hypothesis advanced by Durlach (1963) to explain binaural masking-level difference phenomena. The CMR version of the EC hypothesis suggests that the envelopes of the stimuli in the two auditory channels are subtracted from one another (Buus, 1985; Cohen and Schubert, 1985; Haggard et al., 1985; Hall, 1986). If the envelopes are correlated, following subtraction, only the pure-tone signal will remain and can thus be detected. According to this hypothesis, a CMR effect should be measured regardless of whether the tone is placed exclusively in the dips or exclusively in the peaks of the temporal envelope of the masker-band; following a process of subtraction the residual energy resulting from the tone will remain. Grose and Hall (1989), however, failed to measure a CMR effect if they placed a pure-tone signal in only the peaks of the envelope of the masker-band, but measured a substantial CMR when placing the signal in only the dips of the masker. Thus, this hypothesis also fails to explain CMR. A third hypothesis based on an across-channel comparison suggests that CMR results from detection of a decorrelation of temporal envelopes that were previously correlated. The
envelopes become decorrelated when the pure-tone signal is added to the masker-band (Hall et al., 1984; Cohen and Schubert, 1985, 1987a; Richards, 1987). In support of a decorrelation cue, Richards (1987) has demonstrated that spectrally separated noise-bands with correlated temporal envelopes can be distinguished from spectrally separated noise-bands with uncorrelated envelopes. Hall and Grose (1988), however, measured a significant CMR effect using a signal identical to the masker-band, which should not cause a decorrelation of temporal envelopes. That is, a signal identical to the masker has no effect on the pattern of amplitude fluctuation in the masker-band and thus the envelopes of the masker- and flanker-band remain correlated. Therefore, for this type of stimulus, the observed CMR effect cannot be accounted for by a cue of envelope decorrelation. In addition, McFadden (1986, 1987) failed to measure a CMR effect using a narrow-band noise signal rather than a pure-tone. A noise-band signal will alter the pattern of modulation in an AM masker as readily as a pure-tone signal. Therefore, failure to observe a CMR effect using a noise-band signal also speaks against a CMR hypothesis based on a decorrelation cue.

GROUP 2 HYPOTHESES

(NO ACROSS-CHANNEL COMPARISON)

Many of the hypotheses of the second group (those not based on an
across-channel analysis) suggest that CMR may be explained by within-channel cues (Schooneveldt and Moore, 1987). A CMR effect also occurs, however, if the flanker-band is presented to the ear contralateral to the ear the masker-band and pure-tone signal are presented (Hall et al., 1984; Hall, 1986; Cohen and Schubert, 1987b; Schooneveldt and Moore, 1989a). This finding alone refutes any CMR hypothesis based exclusively on a within-channel cue.

Buus (1985) has proposed an hypothesis based neither on within-channel cues nor on across channel comparisons. According to his model the correlated flanker-band is used by the auditory system to identify the places in the masker-band where the best signal-to-noise ratios exist. In other words, the correlated flanker-band helps the auditory system identify the ideal times to listen for a signal in the masker-band. This information would be unavailable, however, if a pure-tone is added also to the flanker-band. In doing just that, Hall et al. (1988) still measured a substantial CMR effect.

AN HYPOTHESIS BASED ON AUDITORY-GROUPING

A CMR hypothesis that has received very little attention, until recently, is one that suggests that the CMR effect may result from the perceptual grouping of auditory stimuli. Although this possibility was alluded to as early as 1984 by Hall et al., only recently has research surfaced that addresses a CMR hypothesis based on the phenomena of auditory-grouping.
Auditory-grouping refers to ways the auditory system organizes a complex acoustic environment. [The terms *auditory scene analysis* and *stream segregation* have also been used to refer to the perceptual grouping of acoustic energy (see Bregman, 1990)] In general, upon encountering a complex mixture of sounds, the auditory system seeks to organize the mixture into separate perceptual groups (see Handl, 1989). [Auditory, perceptual groups have also been referred to as *streams* or *auditory streams* (see Bregman, 1990)] Usually, all the frequency components of a particular auditory group can be traced back to the same sound source or acoustic object (Bregman and Pinker, 1978). Therefore, it appears that the auditory system tries to group together frequency partials that arise from the same source. The advantage of auditory grouping is that the auditory system can then process each group as a unit, instead of trying to independently process each frequency partial in a complex acoustic mixture.

Sounds that are similar to each other are usually channelled into the same auditory group. Separate sounds may be viewed as similar to each other based on frequency proximity, intensity ratio, a common fundamental, or by synchronous fluctuations in their amplitude or frequency (for a review of acoustic relationships that contribute to auditory grouping see Bregman, 1978, or Moore, 1989). Sounds that are similar and occur simultaneously are often grouped together. When this occurs the sounds are said to be "fused" or "vertically-fused" (Bregman and Campbell, 1971; Dannenbring and Bregman, 1989).
Sound components that are fused evoke a single perceptual image, regardless of the number of sound components in the group. A simultaneous presence of separate sounds, however, is not mandatory for separate sounds to be placed into the same auditory group. In other words, sounds that occur at different moments in time also may be grouped together (Bregman and Campbell, 1971; Bregman and Rudnicky, 1975; Dannenbring and Bregman, 1976; Bregman, 1978). When sounds are grouped in this way, they are said to be "streamed" together or that they are in the same "horizontal-stream." For example, if three high-frequency tones and three low-frequency tones are interlaced temporally (alternate high-frequency and low-frequency tones) and are played repeatedly in a single, rapid series, two horizontal-streams will emerge. That is, after a few cycles, a listener will hear the one series of tones split into two perceptual streams (Bregman and Campbell, 1971). One stream will be composed of only the low tones and the other of only the high tones.

In terms of the CMR effect, vertical-fusion is the type of auditory grouping that may play a substantive role in signal detection improvement. Grose and Hall (1990) gave the first published argument addressing an auditory-grouping hypothesis. They found that CMR was sensitive to signal frequency uncertainty. A disruption of CMR by signal frequency uncertainty suggests an interference in signal detection and not a hindrance to some type of across-channel difference detection. Said another way, this finding suggests that it is the signal itself that is being detected in CMR experiments, and therefore the CMR effect must result
from an enhancement of the pure-tone signal in the correlated noise-band conditions, rather than a detection of dissimilarity between correlated noise-bands. Grose and Hall (1990) suggest that the tone may be enhanced by auditory fusion of the noise-bands when their temporal envelopes are correlated. Fusion of the correlated noise-bands may leave the static pure-tone perceptually isolated from the noise-bands and, thus, "easier" to detect. Therefore, perceptual isolation of the tone when the temporal envelopes of the noise-bands are correlated may be the underpinning of the CMR effect.

Hall and Grose (1990) also provide evidence that supports a CMR hypothesis based on auditory grouping. Using multiple flanker-bands, they explored CMR in conditions wherein some flanker-bands were comodulated with the masker-band, and other flanker-bands were not; those noise-bands that were not comodulated were called deviant-bands. The envelopes of the deviant-bands were either identical to one another (codeviant-bands) or independent of one another (bideviant-bands). Both codeviant- and bideviant-bands reduced the magnitude of the CMR effect; however, the reduction was greatest for bideviant-bands. Adding deviant-bands to the otherwise comodulated noise-band complex may cause the process of perceptual grouping to become more difficult, because more auditory groups must be formed, and this may possibly weaken the perceptual boundaries between auditory groups. If this is so, the perceptual isolation of a static pure-tone from modulated noise-bands is decreased. If the CMR effect is truly a manifestation
of auditory grouping, this decrease in perceptual isolation of the pure-tone should be reflected by a reduction in the magnitude of the effect. In the experiment of Hall and Grose (1990) the auditory system must form one additional perceptual group if the deviant-bands are codeviant, but must form at least two additional perceptual groups if the deviant-bands are bideviant. Therefore, in the bideviant-band condition, perceptual boundaries between auditory groups are weaker and perceptual isolation of the pure-tone signal is lower. This would explain the greater reduction in the CMR effect in bideviant-band conditions than in codeviant-band conditions.

Minimal work has been done to test a CMR hypothesis based on auditory grouping. The purpose of this investigation, therefore, is to determine if a relationship exists between the perceptual organization of sound (i.e., auditory-grouping) and the CMR effect. To examine the possibility of this relationship, a series of three separate, but related experiments is presented. In the first experiment, signal detection thresholds are obtained in both correlated and uncorrelated noise-band conditions. This is done to determine if a CMR effect occurs using the amplitude modulated (AM) noise-bands constructed specifically for this investigation. The purpose of the second experiment is to determine if "strength" of vertical-fusion is greater for AM noise-band pairs with correlated temporal envelopes than for noise-band pairs with uncorrelated temporal envelopes. In experiment three, the temporal envelopes of the noise-bands are correlated in the two conditions that are compared. In one condition,
detection thresholds for a pure-tone are obtained with the correlated flanker-band perceptually "stripped" from vertical-fusion by capturing it into a horizontal-stream. The thresholds obtained in the "flanker-band stripped" conditions are compared with thresholds obtained in conditions wherein the correlated noise-bands remain fused by the auditory system. Therefore, in the third experiment, although the temporal envelopes are always identical within each noise-band pair, the strength of vertical-fusion of the correlated noise-bands varies across the two conditions. This may allow for an examination of the role played by auditory fusion in enhancing detection of a tonal signal in coherently modulated noise-bands.
EXPERIMENT 1

(Comodulation masking release)

INTRODUCTION

This CMR experiment is similar to the CMR experiments completed by others (for a review of experiments in CMR see Hall, 1987); however, it is unique in several ways. For one, the duration of both the AM noise-bands and the pure-tone signal is considerably shorter than those used in most other CMR experiments. For another, only five pairs of "uncorrelated" noise-bands constitute the uncorrelated condition in this experiment (see footnote 1 at end of this experiment) -- as well as in the two experiments that follow. In other studies of CMR that have used digitized noise-bands (i.e., "frozen" noise), the uncorrelated flanker-band has been selected at random from an array of digitized flanker-bands (see McFadden, 1986). Random selection of the flanker-band creates an uncorrelated condition that is similar to an analog generated, uncorrelated condition (the vast majority of CMR experiments have used analog systems). The relationship between the temporal envelopes of spectrally separated noise-bands is continuously changing within an uncorrelated condition that is analog driven. By randomly selecting the flanker-band in a digitally driven system, the relationship between the temporal envelopes of the noise-bands also changes frequently, approximating an analog system --
although specific temporal relationships between envelopes are repeated. In addition, random selection of the uncorrelated flanker-band guards against a learning effect. On the other hand, limiting the uncorrelated condition to only several set relationships between the temporal envelopes of masker-band and flanker-band creates an uncorrelated condition that can be analyzed in greater detail than an ongoing (analog) or completely randomized (analog-like) condition. Studying details of the uncorrelated condition may shed light on the problem of why a pure-tone can be detected at a lower decibel level if the temporal envelope of a flanker-band fluctuates in synchrony with the masker-band than if it does not. For this reason, only five pairs of noise-bands constitute the uncorrelated condition. Because the noise-bands are digitized, the temporal relationship between the noise-bands within each of the five pairs does not change. Finally, this CMR experiment differs from others in that a tracking procedure is used to obtain detection thresholds for the pure-tone signal. All previous studies of CMR have used a method-of-limits procedure in a 2 or 3 alternative forced choice paradigm (e.g., Hall, et. al., 1984; Hall, 1986a, 1986b; Schooneveldt and Moore, 1987, 1989; Hall and Grose, 1988, 1990; McFadden, 1986, 1987).

It is possible that a CMR effect may not occur using the short duration stimuli, relatively "static" uncorrelated condition, and threshold procedure selected for this study. Therefore, the purpose of this first experiment is to
determine if the subjects participating in the experiment demonstrate a CMR effect, and if so, the magnitude of the effect.

METHODS

A. SUBJECTS

Three subjects participated in this experiment (subjects JC, MC, and WJ). The age range of the subjects was from 31 to 54 years. Subjects JC and MC had participated in many other psychoacoustic experiments, and therefore both were highly experienced listeners. Although subject WJ had not participated in any other psychoacoustic experiments prior to this one, she was an Audiologist by profession and was considered a "sophisticated" listener. During this experiment, and the two that follow, subjects JC and MC listened with their left ear; subject WJ listened with her right ear.

All three subjects had pure-tone thresholds of 10 dBHL or less at the octave related frequencies between 250- and 8000-Hz for the ear they listened with during the experiment. Also, when listening with the "test ear," each subject had a word recognition performance score of 100%. (See appendix A for additional audiometric information.)
B. STIMULI

Ten narrow-bands of amplitude modulated (AM) noise were constructed and digitally stored in the computer. Five of the noise-bands were spectrally centered at 1700-Hz. The temporal envelope of each noise-band centered at 1700-Hz was different from the temporal envelopes of the other four noise-bands centered at this frequency. In other words, none of the five envelopes were the same. The other five noise-bands were centered at 2000-Hz. The five different envelopes of these noise-bands were identical to the envelopes of the noise-bands centered at 1700-Hz. That is, for each of the five different noise-band envelopes centered at 1700-Hz, there was a noise-band with the same temporal envelope centered at 2000-Hz. The spectral width of all AM noise-bands was 26 Hz. Duration of the noise-bands was 320-ms, with 10-ms linear rise and fall times.

The five different envelopes are denoted by letters of the alphabet (A, C, D, E, and H). The ten noise-band pairs are represented by these letters also. Within a pair, the first letter denotes the envelope of the noise-band centered at 1700-Hz, and the second letter denotes the envelope of the noise-band centered at 2000-Hz. The correlated noise-band pairs, of course, are identified by the same letter printed twice (e.g. AA and DD). In this investigation, the specific uncorrelated noise-band pairs are CA, DC, ED, HE, and AH. The uncorrelated noise-band pairs were created by simply pairing the envelope of the noise-band centered at 2000-Hz with the noise-band centered at 1700-Hz.
that had the temporal envelope denoted by the next letter of the alphabet used in this investigation (e.g., noise-band centered at 2000-Hz having envelope "C" was paired with noise-band centered at 1700-Hz having envelope "D," written as DC). The noise-band centered at 2000-Hz with envelope "H" was paired with the noise-band centered at 1700-Hz having envelope "A," because the noise-band with envelope "A" was the only one that was not used in the other pairs.

The signal detected in this experiment was a 2000-Hz pure-tone. This signal frequency was selected because the magnitude of the CMR effect has been found to be greater in the higher frequencies (2000 Hz and above) [Cohen and Schubert, 1987b; Schooneveldt and Moore, 1987]. The duration of the pure-tone was 140-ms, with a 10-ms linear rise and fall time. It was spectrally centered in the noise-bands centered at 2000-Hz, and temporally centered within the 320-ms duration of all noise-bands. See figure 1.

**CONSTRUCTION OF AM NOISE-BANDS**

The noise-band stimuli were constructed by a three step process and digitally stored in an 80286-based microcomputer (Northgate, 286-12). In step one, 320-ms samples of low-pass noise were created by passing a wide-band of random noise through a low-pass filter (Wavetek, 752A, Brickwall, 115 dB/oct skirts) with a high frequency cut-off of 13-Hz. These samples of low-pass noise were digitally stored following analog-to-digital conversion at a sampling rate of
FIGURE 1.
Temporal aspects of noise-bands and pure-tone signal. Both the noise-bands and pure-tone have a linear, 10-ms rise and fall time. The tone is temporally centered in the noise-bands. A flanker-band with a correlated temporal envelope is shown.
10000-Hz. (These samples served as modulators.) In step two, each sample of noise was digitally multiplied by the sine function:

$$\sin \left[ 2\pi f_c(i) \right];$$

$$i = 0, 1, 2, \ldots$$ at discrete values of \(t\) corresponding to \(1/f_s\), the sampling frequency of 10000-Hz.

The \(f_c\) determined the carrier frequency or center frequency of the resulting AM noise-band. The product of each multiplication process was a narrow-band of noise with a 26 Hz spectral width, amplitude-modulated at the fluctuation rate of the sample of low-pass noise, and centered at the frequency of the sinusoidal carrier (i.e., \(f_c\)). In step three, the digital version of each was adjusted so the peak-sound-pressure-level (pSPL) of all noise-bands differed by less than .25 dB, except for the two noise-bands with the temporal envelope denoted by the letter "E." The pSPL of the noise-bands with envelope E was about 3 dB greater than all the other AM noise-bands. (See footnote 2 at end of this experiment.) A visual representation of each AM noise-band is shown on the next page in figure 2.
AMPLITUDE MODULATED NOISE-BANDS

Figure 1 ENVELOPE A; CARRIER = 1700 Hz

Figure 2 ENVELOPE C; CARRIER = 1700 Hz

Figure 3 ENVELOPE D; CARRIER = 1700 Hz

Figure 4 ENVELOPE E; CARRIER = 1700 Hz

Figure 5 ENVELOPE H; CARRIER = 1700 Hz

Figure 6 ENVELOPE A; CARRIER = 2000 Hz

Figure 7 ENVELOPE C; CARRIER = 2000 Hz

Figure 8 ENVELOPE D; CARRIER = 2000 Hz

Figure 9 ENVELOPE E; CARRIER = 2000 Hz

Figure 10 ENVELOPE H; CARRIER = 2000 Hz

FIGURE 2.

Center frequency is 1700-Hz for the AM noise-bands in the left column; they served as flanker-bands. Center frequency is 2000-Hz for the AM noise-bands in the right column; they served as masker-bands. All noise-bands are 320-ms in duration with 10-ms linear rise and fall times, and have a spectral width of 26-Hz.
CALIBRATION

Calibration of the pSPL of the noise-band stimuli and the dBsPL of the pure-tone signal was completed prior to running subjects each day. Because of the short duration of the stimuli, they were calibrated by, first, splitting the final output and sending it to an earphone and an oscilloscope in parallel. A measure of the maximum voltage (peak-to-peak) level displayed on the CRT of the oscilloscope was then recorded for each noise-band, as well as the "shaped tone." After voltage levels were recorded, a continuous sinusoid was then sent, in the same manner, to both the oscilloscope and earphone. The voltage (peak-to-peak) level of the sinusoid was adjusted to match that recorded for each noise-band and the shaped pure-tone. For each noise-band, after the voltage level was matched, the acoustic output of the earphone was measured in dBsPL using a flat-plate coupler, B & K half-inch microphone (Model 4134) and B & K measuring amplifier (Type 2609). This dBsPL value plus 3 dB represented the pSPL of a matched noise-band stimulus. Similarly, the dBsPL of a continuous pure-tone matched to the peak-to-peak voltage of the shaped tone was recorded as the dBsPL of the shaped pure-tone signal.

C. INSTRUMENTATION (See figure 3.)

A digitally stored ramp was D/A converted by a M108 module of the Modular Instruments (MI²) multi-function unit. The ramp was 140-ms in
duration and temporally centered in a 320-ms file. The ramp was routed to a locally built multiplier, where it was multiplied with a 2000-Hz sinusoid. The sinusoid was generated by a Wavetek (Model 148A) signal generator. The output of the multiplier (2000-Hz pure-tone shaped by 140-ms ramp) was patched into a dual attenuator module (M208) of the MI² unit. The output of the dual attenuator was sent to mixer 2, where it was mixed with the AM noise-bands.

The digitally stored noise-bands were converted from digital-to-analog form by a D/A module (M108) of the MI². The output (noise-bands) of each of two ports of the D/A module was fed into a locally built mixer (mixer 1). The output of mixer 1 was routed to an Hewlett Packard power attenuator (Model H-P 350D), via a resistive matching network. The output of the attenuator was impedance-matched and fed into mixer 2 of the locally built mixer, where they were mixed with the pure-tone signal.

The output of mixer 2 went to a Wavetek/Rockland brickwall filter (Model 752A), where the composite stimulus was low-pass filtered at 2300-Hz to eliminate high frequency energy caused by D/A conversion. The output of the filter was fed to a power amplifier (SAE; Model A202), then to a step attenuator (Shalco; Model T-320-B), and, finally, from the step attenuator to an earphone mounted in a circumaural cushion (Sennheiser; Model HD430).
Instrumentation diagram for CMR experiment. See written description.
D. EXPERIMENTAL PROCEDURE

Signal detection thresholds for a 2000-Hz pure-tone centered in AM narrow-band maskers were obtained by a tracking procedure. Subjects controlled attenuation of the pure-tone signal by two hand-held, button type switches. Pressing the button down of one switch caused an increase in attenuation of the tone, pressing the button down of the other switch caused a decrease in attenuation.

The subjects were instructed to press the button that increased the attenuation of the tone when the tone became audible, and to press the button that decreased the attenuation of the tone when the tone became inaudible. Attenuation incremented in .375 dB steps. [The procedure was much like that used in Bekesy-audiometry (see Bekesy, 1947 or Reger, 1952).]

Detection thresholds for the pure-tone were obtained in the five pairs of correlated noise-bands and the five pairs of uncorrelated noise-bands. Five replicates for threshold were obtained for each of the ten pairs of noise-bands. Threshold was obtained by determining the average level of attenuation, following the third reversal.

Except for the noise-bands modulated by envelope E, each noise-band was presented at 61 dB pSPL. The two noise-bands with the E envelope were presented at 64 dB pSPL -- 3 dB higher than the others. There was 480-ms between the end of one noise-band pair and the beginning of the next. There were 167 occurrences of the noise-bands and pure-tone signal during each
experimental run; therefore, each run lasted for about 2 minutes and 15 seconds. A typical threshold track is shown in figure 4.

RESULTS

Significant differences were found (based on a $3 \times 10 \times 5$, within-subject, repeated measures ANOVA, cond $[F(9,18) = 5.54] p<.01$) across the thresholds obtained in the ten separate pairs of AM noise-bands (appendix B). A significant difference was not found for replicates (reps $[F(4,8) = .56]$ NS), nor was a significant interaction noted between replicates and noise-band pairs (cond * reps $[F(36,72) = .91]$ NS); both suggested the absence of a learning effect. Thresholds in the correlated condition differed significantly from those in the uncorrelated condition ($[F(1,18) = 8.44] p < .01$; see contrast statement in appendix B). The overall CMR effect was 2.8 dB (see figure 5). That is, pooling data across subjects, the mean threshold was 2.8 dB lower in the correlated noise-bands than in the uncorrelated noise-bands.

Duncan’s, post-hoc, multiple range test revealed significant differences between thresholds obtained in some of the correlated noise-band pairs and uncorrelated pairs, but not all (appendix C). The comparisons of particular interest were those between the correlated envelope conditions and uncorrelated envelope conditions, wherein the envelopes of the masker-bands
A SAMPLE TRACING OF SUBJECT MC
TRACKING THE 2000-Hz PURE-TONE SIGNAL

FIGURE 4.

The abscissa shows the 167 intervals of an experimental trial. The intervals on the ordinate are in .375 dB units of attenuation. Only those levels of attenuation occurring after the third reversal (dark vertical line) were included in the analysis for the mean level of attenuation.
FIGURE 5.

The bar labeled "cor" represents the mean threshold level for the pure-tone signal in the correlated noise-band condition. The bar labeled "uncor" shows the mean threshold in the uncorrelated condition. Bar "cmr" shows the difference between thresholds in the uncorrelated and correlated conditions; i.e., the magnitude of the CMR effect.
were identical across the two conditions. Only two of these five comparisons differed at a significant level, according to Duncan’s test (protection level = .05).

In the two comparisons that differed significantly (CA - AA and ED - DD), the thresholds obtained in the uncorrelated condition were higher than those in the correlated condition; i.e., a CMR effect.

Figure 6 shows the magnitude of the CMR effect at each of the five specific correlated/uncorrelated comparisons of primary interest. The figure demonstrates clearly why the two comparisons (CA - AA and ED - DD) were found to differ significantly by Duncan’s multiple range test.

**DISCUSSION**

Although the overall CMR effect was statistically significant, threshold differences varied considerably across the five correlated/uncorrelated comparisons wherein the masker-band within a comparison was the same in the correlated and uncorrelated pair. (Refer to figure 6.) A large CMR effect occurred for only two of the five comparisons. Clearly, it was the large CMR effect at these two comparisons that caused the overall CMR effect to be statistically significant.

If a relationship exists between CMR and the perceptual organization of sound, then, perhaps, perceptual organization of the correlated and uncorrelated noise-band pairs did not differ substantially in the comparisons for
The magnitude of the CMR effect at each of the comparisons wherein the temporal envelope of the masker-band is held constant.
which a significant CMR effect was absent. Said another way, perhaps the degree of auditory fusion was nearly equivalent between the correlated and uncorrelated noise-band pairs that were compared and did not elicit a CMR effect. If so, then the perceptual organization of the uncorrelated noise-bands must have approached a high degree of auditory fusion, because, presumably, all correlated noise-band pairs were strongly fused.

There were a couple of acoustic characteristics shared by all noise-bands that may have influenced the auditory system to fuse (or partially fuse) some of the uncorrelated pairs of noise-bands. First, the noise-bands of each pair were synchronous in onset and offset. In the "real world," it is highly unlikely that two spectrally separated sounds will begin and end at the same time, and not originate from the same acoustic object. Therefore, it would be sensible for the auditory system to assign sounds that begin and end simultaneously to the same perceptual group (see Bregman, 1990). Synchrony in onset and offset, however, cannot explain why a CMR effect occurred for some of the comparisons but not others, because all pairs of noise-bands were synchronous at both onset and offset. A second acoustic characteristic of the noise-bands that may have influenced the auditory system to fuse some of the uncorrelated pairs was a general similarity in the temporal envelopes of the noise-bands. Although the envelopes used in this study were not identical, the rate of amplitude fluctuation was about the same across all five envelopes. Given the short signal duration and the similar fluctuation-rates, some of the
uncorrelated noise-band pairs might have been sufficiently correlated to elicit a degree of auditory fusion. Perhaps the envelopes of the uncorrelated noise-band pairs were more similar in the comparisons where a CMR effect was not observed, than in the comparisons where a CMR effect was observed.

**TEMPORAL ENVELOPE CORRELATION OF "UNCORRELATED" NOISE-BAND PAIRS AND ITS POSSIBLE RELATIONSHIP TO CMR**

To investigate a possible relationship between envelope similarity, auditory fusion, and the CMR effect, a computer program was developed to determine the degree of correlation of the envelopes of the uncorrelated noise-band pairs. Figure 7 shows correlation coefficients for each of the five uncorrelated noise-band pairs; the correlation coefficient of the correlated noise-band pairs was 1. The pairs that were more highly uncorrelated (correlation coefficients nearest to zero) were the same uncorrelated noise-band pairs used in the comparisons in which a significant CMR effect was found. Those uncorrelated noise-band pairs having the largest absolute value of correlation -- and thus indicating a temporal relationship between the envelopes of the two noise-bands within the pairs -- were the pairs used in the comparisons wherein a substantive CMR effect was not found. Figure 8 shows the magnitude of the CMR effect for each of the five comparisons plotted against the correlation
The coefficients represent the degree to which the envelopes of the two noise-bands are correlated in each uncorrelated (pseudo-uncorrelated) noise-band pair.
FIGURE 8.

Magnitude of the CMR effect for each of the five noise-band pair comparisons as a function of the degree of correlation of the "uncorrelated" noise-band pair involved in each comparison. Bars indicate standard deviation of the means across subjects.
coefficient found for the uncorrelated noise-band pair used in each of the comparisons.

The finding that the uncorrelated noise-band pairs with the lower correlation coefficients (more highly uncorrelated temporal envelopes) were those used in the comparisons where CMR effects were observed, may be interpreted as consistent with a CMR hypothesis based on auditory-grouping. The greater the difference between the envelopes of two co-occurring noise-bands, the less likely they will be fused by the auditory system -- according to the perceptual heuristic "common fate" (Bregman et. al., 1985; Bregman and Pinker, 1978; McAdams, 1982). The weaker the fusion of an uncorrelated pair of noise-bands, the larger the difference between the perceptual organization of the uncorrelated noise-bands and the perceptual organization of the correlated noise-bands. According to the auditory-grouping hypothesis of CMR, the greater the contrast between the perceptual organization of the correlated noise-bands and the uncorrelated noise-bands, the greater the likelihood that the pure-tone signal will be detected at a lower decibel level in the correlated condition than in the uncorrelated condition. By the same token, as perceptual organization of the correlated and uncorrelated noise-band conditions become more similar, the less likely thresholds will differ across the two conditions. This line of thought seems to be supported by the relationship found between CMR and degree of correlation of the uncorrelated noise-band pairs.
Signal Detection Thresholds in Correlated Noise-Band Pairs, and CMR

It can be seen in figure 9 that the two lowest thresholds in the correlated condition are in the correlated noise-band pairs (AA and DD) associated with the largest CMR effects. Low thresholds in these particular correlated pairs may be related to the deep "valleys" (or minimal energy) present in the temporal envelopes of these noise-bands during the time interval the pure-tone signal occurs (see figure 2). This is suggested because there is evidence that a signal is detected primarily in the minima of the envelope of a masker that fluctuates slowly in amplitude (Fastl, 1975; Buus, 1985; Mott and Feth, 1986). Although the deep valleys, or large "dips," may explain the reason the tone is detected at a lower level for these particular correlated noise-band pairs, they do not explain the cause of the large CMR effects associated with these noise-bands. However, large CMR effects associated with noise-bands having large dips in their envelopes during the interval the tone occurs, are consistent with the findings of Grose and Hall (1989). They found a substantial CMR effect if the pure-tone occupies the dips of the temporal envelope of a masker-band, but no CMR effect if the signal occupies only the peaks of the masker-band envelope. The findings of Grose and Hall (1989), and the finding here of large CMR effects for comparisons wherein the masker-band has minimal energy present during the interval the tone occurs, are consistent with the "dip-listening" or "listening in the valleys" hypothesis proposed by Buus (1985).
MEAN_THRESHOLDS_IN_EACH_CORRELATED_NOISE-BAND_PAIR_PLOTTED_AGAINST_THE_CMR_EFFECT_IN_WHICH_EACH_CORRELATED_PAIR_WAS_INVOLVED

FIGURE 9

Mean threshold for the pure-tone in each of the five correlated noise-band pairs is plotted on the abscissa. The magnitude of the CMR effect for the comparison each correlated noise-band pair was involved in, is plotted on the ordinate. Bars indicate ± standard deviation calculated across subject means.
Footnote 1: The five pairs of noise-bands constructed by pairing noise-bands with temporal envelopes that are not identical, constitute the condition referred to as the "uncorrelated condition." The term "uncorrelated condition" implies that a temporal relationship does not exist between the envelopes of the noise-bands within the pairs. This implication is not accurate, however. Because of the short duration of the noise-band pairs, a degree of correlation is present between the envelopes within the pairs. Therefore, these pairs are not completely uncorrelated, but rather pseudo-uncorrelated. Nevertheless, the term "uncorrelated" is used throughout this paper to refer to the pseudo-uncorrelated pairs, so that the basic distinction is clear between the two conditions, and to remain consistent with the terminology used in other studies of comodulation masking release.

Footnote 2: Initially, the AM noise-bands were adjusted in digital form to equivalent RMS values. This process created differences in crest factors across the various noise-bands, which was unwanted. It has been suggested that a signal in AM noise is detected in the minima of the envelope of the masker (Fastl, 1975; Buus, 1985; Mott and Feth, 1986). If this is so, the crests, or peaks, in the envelope of AM noise may contribute to masking the signal via
forward and backward masking (see Schooneveldt and Moore, 1987).

Variations in amplitude of the peaks should produce different levels of effective masking. Therefore, in this investigation, it was decided that crest factors should be held nearly constant across most of the noise-bands. In addition, equating crest factors facilitated calibration.

The pSPL of noise-bands with envelope "E" was found to be about 3 dB greater than all other noise-bands due to an initial calibration error.
EXPERIMENT 2

(Vertical-fusion strength of correlated and uncorrelated noise-band pairs)

INTRODUCTION

The fundamental claim of the auditory-grouping hypothesis of CMR is that strength of vertical-fusion is greater for spectrally separated AM noise-bands that fluctuate in synchrony, than for spectrally separated AM noise-bands that fluctuate independent of each other. The main purpose of this experiment, therefore, is to determine if strength of fusion is greater for the correlated noise-band pairs (used in these experiments) than for the uncorrelated noise-band pairs. Prompted by the findings of experiment 1, possible differences in the degree of auditory fusion across the five uncorrelated noise-band pairs is also investigated.

METHODS

A. SUBJECTS

The three subjects that participated in experiment 1 participated in this experiment also. Two additional subjects (subjects LM and PR) participated in this experiment, for a total of five subjects. Including these two subjects
lowered the floor of the age range to 26 years. Pure-tone thresholds for subject LM were less than 10 dBHL at the octave related frequencies between 250-Hz and 8000-Hz for the ear subject LM listened with during the experiment. Using this ear, subject LM had a word recognition performance score of 100%. Subject PR had pure-tone thresholds of less than 15 dBHL at most audiometric frequencies for the ear he listened with during this experiment, but hearing sensitivity for this ear dropped to 20 dBHL at 4000-Hz and 30 dBHL at 6000-Hz, before rising to 10 dBHL at 8000-Hz. Subject PR had a word recognition performance score of 96% in the ear he used during the experiment (see appendix A). Both PR and LM listened with their right ear during the experiment. It may also be important to note that subject PR is the author of this work.

B. STIMULI AND CALIBRATION

The AM noise-band stimuli and method of calibration were identical to the same described in experiment 1. Only the ten specific pairs of noise-bands used in experiment 1 were used in this experiment.

C. INSTRUMENTATION (See figure 10.)

The digitally stored noise-bands were converted from digital to analog form by a D/A module (M108) of a Modular Instruments (MI) multi-function unit. One output port (2000-Hz noise-band) of the D/A module was routed to
an attenuator module (M208) of the MI² that toggled between maximum attenuation and minimum attenuation, acting as a switch. The other output (1700-Hz noise-band) of the D/A module was routed to a second attenuator module of the MI²; the attenuation level of this module remained constant at minimum attenuation. The output of both attenuators were fed into a locally built mixer. The output of the mixer was then routed to an Hewlett Packard power attenuator (Model H-P 350D) via a resistance matching network. Output of the attenuator was sent through a Wavetek/Rockland brickwall filter (Model 752A), where the composite stimulus was low-pass filtered at 2300-Hz to eliminate high frequency energy caused by D/A conversion. From here the signal went to a power amplifier (SAE; Model A202), a step attenuator (Shallco; T-320-B), and, finally, from the step attenuator to an earphone mounted in a circumaural cushion (Sennheiser; Model HD430).

Two push-button type switches were plugged into a MI² module that interfaced directly with the digital input/output module of the MI², and were used to vary the inter-stimulus-intervals.
INSTRUMENTATION DIAGRAM FOR

EXPERIMENT 2

FIGURE 10.
Instrumentation diagram for investigating strength of vertical-fusion for correlated and uncorrelated noise-band pairs.
D. EXPERIMENTAL PROCEDURE

BACKGROUND AND RATIONALE

In an attempt to quantify "strength" of auditory fusion, a procedure was adopted based on the work of van Noorden (1975), and Bregman and his colleagues (Bregman and Pinker, 1978; Dannenbring and Bregman, 1978; Bregman et al., 1985).

Under normal conditions, tones that are harmonically related will fuse into a single perceptual image. Bregman and his colleagues, however, have shown that one component (target-tone) of a harmonic complex can be stripped from fusion and perceptually placed into a horizontal-stream with "captor-tones" of the same (or nearly the same) frequency. To capture a partial into a horizontal-stream, the complex comprising the target-tone (partial) was alternated at a rapid rate with a sinusoid that was identical in frequency, or close in frequency, to the frequency of the target-tone. To determine if the target-tone was being pulled out of vertical-fusion and into a horizontal-stream, a procedure introduced by van Noorden (1975) was used. This procedure required subjects to judge the rate they heard a tone at the frequency of the target- and captor-tone. If this tone was heard at a rate that was twice that of the other component(s), then the target-tone was presumed extracted from the complex and placed into a horizontal-stream with the captor-tones. If, on the other
hand, the tone was heard at the same rate as the complex (i.e., the multi-component signal), and as alternating with the complex, then complete fusion of the target-tone with the other members of the complex must have been sustained.

Perception of the target-tone was not necessarily exclusively in vertical-fusion with the complex or in horizontal-streaming with the captor-tone, however. That is, the two perceptions were not binary. The prominence of a perception of horizontal-streaming varied according to acoustic relationships between the target-tone and the other partials of the complex. For example, listeners judged the perception of horizontal-streaming to be greater if the onset of the target-tone preceded the onset of the other partials (Bregman and Pinker, 1978; Bregman et al., 1985). Also, the percept of horizontal-streaming was usually more prominent if the decibel level of the target-tone was greater relative to the other partials. Listeners were able to rate the prominence of horizontal-streaming on a qualitative scale (Bregman and Pinker, 1978; Dannenbring and Bregman, 1978). They were also able to rate "decomposition" of the complex. A high decomposition rating suggested weak vertical-fusion; a low decomposition rating suggested strong vertical-fusion. High decomposition ratings coincided with high prominence of horizontal-streaming.

Using van Noorden's capturing procedure, Bregman et al. (1985) studied the influence of amplitude modulation (AM) on auditory-grouping. They found that a complex of sinusoidally-amplitude-modulated (SAM) tones were strongly
vertically-fused into one auditory image if the applied frequency and phase of SAM was the same for each tone in the complex. If the phase or frequency of SAM of a target-tone was altered relative to the other partials of the complex, fusion was weakened. Therefore, correlated temporal envelopes enhanced fusion strength of a complex.

In the studies described briefly above, the duration of the silent period between the end of one stimulus interval and the beginning of another was constant across the stimulus conditions compared. The listeners reported qualitative judgements concerning the prominence of the percepts horizontal-streaming and vertical-fusion.

Qualitative reports are often quite telling in the realm of perception. For this investigation, however, a quantitative measure of "strength" of vertical-fusion was desired. From the works of van Noorden and Bregman (discussed briefly above), it was reasoned that a shorter silent period between stimulus intervals may be required to capture an acoustic component into a horizontal-stream if the temporal envelope of that component fluctuates in synchrony with other components of an acoustic complex than if the envelope of that component does not fluctuate in synchrony with the envelopes of the other components. This reasoning was the basis of the procedure used in this experiment to investigate the strength of vertical-fusion of correlated and uncorrelated noise-band pairs.
PROCEDURE

In this experiment, the subjects heard a captor-band (AM noise-band centered at 1700-Hz) presented alone alternate with a masker-band and target-band presented together (bands centered at 2000-Hz and 1700-Hz, respectively). The temporal envelope of the target-band was always identical to the captor-band. The temporal envelopes of the noise-band pairs (masker-band and target-band) were either correlated or uncorrelated.

The subjects sat in a sound treated suite wearing the earphones. Each noise-band was presented at 65 pSPL, except for the two noise-bands modulated by the E envelope which were presented at 68 dBSPL. The stimuli were presented monaurally at an initial interstimulus-interval (ISI) of 480-ms between noise-band pairs and captor-bands.

The task of each subject was to adjust the duration of the silent period, or ISI, between the noise-band pairs and the captor-bands until the predominant percept was horizontal-streaming of target-bands with captor-bands. The subjects decreased the duration of the ISI in 5 ms intervals, by pressing down a hand-held button type switch, until the target-band of the noise-band pair was "captured" into a horizontal-stream with the captor-bands; that is, until the predominant percept became one of horizontal-streaming. (See footnote 3 at end of this experiment for discussion of why subjects were not allowed to increase the ISI.) When a listener judged horizontal-streaming to be
the predominant perception, they ended the experimental trial by pressing the two hand-held buttons simultaneously. The ISI at exit was used as an estimate of the temporal "nearness" of captor-band to target-band necessary to strip the target-band from vertical-fusion. Shorter ISIs were interpreted as suggesting greater difficulty in capturing the target-band into a horizontal-stream, thereby suggesting stronger vertical-fusion of target-band with the co-occurring noise-band. By the same token, longer ISIs were interpreted as suggesting weaker auditory fusion of the two co-occurring noise-bands. See figures 11, 12, and 13.
PRESENTATION SEQUENCE OF STIMULI FOR EXPERIMENT 2

FIGURE 11.

Above shows captor-band alternating with a noise-band pair, which includes the target-band. Subjects decreased the duration of the ISI until the target-band was captured into a horizontal-stream with the captor-band.
FIGURE 12.

Above is a visual illustration of the auditory percept of vertical-fusion for the co-occurring noise-bands having correlated temporal envelopes.
FIGURE 13.

Above is a visual representation of the auditory percept horizontal-streaming for target-bands and captor-bands.
RESULTS

The interstimulus-interval (ISI) necessary to capture a target-band into an horizontal-stream differed at a significant level across the ten pairs of noise-bands (based on a 5 x 10 x 8, within-subject, repeated measures ANOVA, cond [ F(9,36) = 3.84] p < .01; appendix D). A level of statistical significance was not found across replications (reps [F(7,28) = .19] NS), nor was a significant interaction noted between replicates and noise-band pairs (cond * reps [ F(63,252) = .84] NS); suggesting the absence of a learning effect. Also, the ISI obtained for the correlated noise-band condition differed significantly from the ISI obtained for the uncorrelated noise-band condition ( [ F(1,252) = 59.61] p < .01; see contrast statement in appendix D and figure 14). The means for ISI in all five uncorrelated noise-band pairs were longer than all five means in the correlated noise-band pairs (See figure 15).

DISCUSSION

The results of this experiment support the hypothesis that strength of vertical-fusion is greater for spectrally separated noise-bands if the modulation applied to them is coherent rather than independent. This is in agreement with the basic premise of the auditory-grouping hypothesis of CMR.
FIGURE 14.

The bar labeled "cor" represents the mean ISI necessary to capture the target-bands into a horizontal-stream for the noise-band pairs with correlated temporal envelopes. The bar labeled "uncor" represents the same for the uncorrelated noise-bands. Bars indicate standard deviation of data across subjects.
FIGURE 15.
Above shows the mean ISI necessary to capture the target-band into a horizontal-stream for each pair of noise-bands. Bars indicate standard deviation of data across subjects.
As was the case for experiment one, the findings of particular interest are those concerned with the uncorrelated noise-band pairs. The ISI required to capture a target-band into a horizontal-stream differed across the five uncorrelated pairs (refer to figure 15). Presumably, the more difficult it is to capture a component of a complex into a horizontal-stream, the greater the strength of vertical-fusion of the complex. Therefore, vertical fusion strength is believed greater between the noise-bands constituting the pairs that require shorter ISIs to elicit a percept of horizontal-streaming, than between the noise-bands constituting the pairs for which longer ISIs are sufficient to elicit a percept of horizontal-streaming. In short, the ISI data suggest that the strength of vertical-fusion is greater for some of the uncorrelated noise-band pairs than it is for others.

**INTERSTIMULUS-INTERVALS AND CORRELATION COEFFICIENTS**

It was suggested in the discussion section following experiment one, that the degree of correlation of the temporal envelopes within an uncorrelated noise-band pair may be closely related to the strength of fusion of the pair. Therefore, two indicators of strength of vertical-fusion have been proposed: 1) the ISI duration to capture a target-band into an horizontal-stream, and 2) the coefficient of correlation of temporal envelopes within a co-occurring pair of noise-bands.
In figure 16 the correlation coefficients of the five uncorrelated noise-band pairs are plotted against the ISIs necessary for horizontal-streaming to occur in the uncorrelated noise-band pairs. The line-chart shows the uncorrelated noise-band pair with the lowest correlation coefficient (ED) to be the one where the target-band was easiest to capture into a horizontal-stream ("easiest" suggested by the longest duration ISI). The line-chart also shows a tendency for capturing a target-band to become more difficult as a temporal relationship emerges between the envelopes within a noise-band pair. This tends to suggest that the temporal envelopes of spectrally separated noise-bands do not have to be identical before the auditory system begins to fuse them. Even small relationships between temporal envelopes of co-occurring noise-bands may contribute to their fusion.

**INTERSTIMULUS-INTervals AND CMR**

According to the auditory-grouping hypothesis of CMR, a pure-tone signal can be detected at a lower decibel level in correlated noise-bands than in uncorrelated noise-bands, because the correlated noise-bands are fused by the auditory system and leave the pure-tone perceptually isolated. The hypothesis suggests that uncorrelated noise-bands are not fused, the pure-tone is not isolated perceptually, and, thus, a higher decibel level is required for the tone to be detected. If the degree of vertical-fusion of noise-bands within an
uncorrelated condition approaches that of noise-bands in a correlated condition, however, the decibel level necessary to make the tone detectable should be equivalent across the two conditions, according to the auditory-grouping hypothesis of CMR. In other words, a CMR effect should not occur if the uncorrelated noise-bands are fused by the auditory system.

The line-chart of figure 17 shows a tendency, albeit a weak one, for the CMR effect to be largest for each subject in those comparisons comprising the uncorrelated noise-band pairs for which longer ISIs were sufficient to capture one noise-band of the pair into a horizontal-stream. The chart also demonstrates that the uncorrelated noise-bands with the shortest ISIs in this experiment were involved in the comparisons that offered the smallest or an absent CMR effect in experiment one. As in the correlation coefficient comparisons, the ISI data suggests that the noise-band pair ED should be the one for which fusion strength is weakest, and therefore be in the comparison that elicits the largest CMR effect. This is not the case, however. The comparison that includes the uncorrelated pair ED demonstrates the second largest CMR effect. The largest CMR effect occurred for the comparison comprising the uncorrelated noise-band CA. The ISI data and the correlation coefficient data suggest that the fusion strength of CA should be greater than that of ED, and, therefore, the CMR effect should be smaller in the comparison involving CA than in the comparison involving ED.
FIGURE 16.

In the above, the ISI needed to capture the target-bands from the uncorrelated noise-band pairs are plotted against the correlation coefficients of the uncorrelated noise-band pairs.
FIGURE 17.

In the above, the ISI needed to capture the target-bands from the uncorrelated noise-band pairs are plotted against the magnitude of the CMR effect involving the same uncorrelated noise-bands.
A longer ISI may have been sufficient to capture the noise-band with envelope E of the ED pair, because the pSPL of this noise-band was 3 dB higher than its partner, which was modulated by envelope D. Increasing the level of a target-component relative to the other components of the complex facilitates horizontal-streaming of the target-component with a captor-component (van Noorden, 1975; Bregman and Pinker, 1978; Dannenbring and Bregman, 1978; Bregman et al., 1985). Therefore, a longer ISI may have been sufficient to capture the "E noise-band" because its decibel level was greater than that of the other noise-band of the pair, and not because strength of fusion was weakest for the ED pair. Lower fusion strength for the ED pair than for the CA pair would be consistent with the larger CMR effect found in the comparison that involves the CA uncorrelated noise-band pair.
Footnote 3: Initially, the subjects were allowed to both increase and decrease the duration of the ISI until they believed the percept to be predominately horizontal-streaming. The variability in the data was extremely high, however, when allowing the subjects to adjust the ISI in both directions. It was believed that there were at least two contributors to the high variability in the initial data.

One cause of the highly variable data may have been related to the *hysteresis of perceptual organization* (Bregman, 1990). In short, the hysteresis of perceptual organization appears to result from a listeners tendency to "hold-on" to perceptions that have already been formed. (For a more elaborate discussion of the hysteresis of perceptual organization see Bregman, 1990.) In this study, if the initial perception was not horizontal-streaming (long duration ISI), the ISI was adjusted to very short durations before horizontal-streaming was perceived. On the other hand, if the initial perception was horizontal-streaming (short ISI) the ISI was adjusted well past the point where the percept became horizontal-streaming, before the duration of the ISI was sufficiently long for horizontal-streaming to no longer be the predominant percept. Therefore, the ISI required for the percept to become one of horizontal streaming, and the ISI required for the percept of horizontal-streaming to cease, were frequently in disagreement.
Also, the initial data may have been highly variable because of the ability of listeners to "shift" from a synthetic mode of listening to an analytic mode of listening. According to McAdams (1983) the default mode is synthetic, and is when co-occurring sounds are perceptually fused. Analytic listening, in contrast, is perceptually parsing an acoustic complex into its separate components and therefore "hearing-out" each component of the complex. Within certain limits, a listener can switch from listening synthetically to listening analytically. In this study, if a subject was listening in an analytic mode, horizontal-streaming was more likely to occur than if the subject was listening in a synthetic mode. Therefore, subjects could often change the predominant perception just by switching from one listening mode to the other.

Because of the hysteresis of perceptual organization and the apparent ability of subjects to switch from synthetic to analytic listening, the initial data were discarded. To obviate the problem of perceptual hysteresis, the subjects were allowed to adjust the duration of the ISI from long to short only. In hope of minimizing problems caused by switching listening modes, each subject was instructed to avoid listening in an analytic mode. In other words, they were told to try and hear the noise-band pair as a whole instead of listening for its individual components.
EXPERIMENT 3

(Detection thresholds in strongly fused and weakly fused correlated noise-band pairs.)

INTRODUCTION

The intent of this experiment is to determine if degree of vertical-fusion of flanker-band and masker-band enhances detection of a pure-tone signal. To examine this, the ideal experiment would manipulate strength of vertical-fusion of masker-band and flanker-band, while leaving all other variables untouched. In other words, degree of fusion of correlated noise-bands would have to be altered without changing the acoustics of the noise-bands. The design of this experiment attempts to create this ideal situation.

Thresholds are obtained only in noise-band pairs with correlated temporal envelopes, but the degree of vertical-fusion between masker-band and flanker-band constituting the pairs is manipulated. Thresholds for the pure-tone in a condition where strength of vertical-fusion is presumed high, are compared to thresholds in a condition where strength of vertical-fusion is presumed low. If thresholds are found to be worse in the weakly fused condition than in the strongly fused condition, support for auditory-grouping as a contributor to the CMR effect is suggested. On the other hand, if a difference does not occur
between thresholds in the strongly fused and weakly fused condition, no support for the auditory-grouping hypothesis of CMR is gained.

METHODS

A. SUBJECTS

The three subjects that participated in experiment 1 (JC, MC, and WJ) were the only three subjects that participated in this experiment. Subjects PR and LM did not participate.

B. STIMULI AND CALIBRATION

The same ten noise-bands used in the first two experiments were used in this experiment also. Only the five correlated noise-band pairs were used, however. As in experiment 1, the signal detected was a 2000-Hz pure-tone shaped by a 140-ms ramp with a 10-ms rise and fall time. Calibration was accomplished in the same manner as in experiment 1 and 2.

C. INSTRUMENTATION

PART A (SEE FIGURE 18.)

The digitally stored noise-bands and ramp were converted from digital to analog form by D/A modules (M108) of a Modular Instruments (MIP), multi-
FIGURE 18.

Instrumentation diagram for determining signal detection thresholds in the correlated noise-band pairs, with the flanker-band captured into a horizontal-stream.
function unit. An output port (1700-Hz noise-band) of one of the D/A modules was routed to a attenuator module (M208) of the MI² that toggled from maximum attenuation to minimum attenuation, acting as a switch. The other output (2000-Hz noise-band) of this same D/A module was routed to a second attenuator module of the MI²; the attenuation level of this channel of this module was constant. One of the output ports (140-ms ramp) of the second D/A module was fed into a locally built multiplier, where it was multiplied by a 2000-Hz sinusoid. The 2000-Hz sinusoid was generated by a Wavetek (model 148A) signal generator. The output of the multiplier (pure-tone shaped by ramp) was first routed to the attenuator of the MI² used as a switch, and then to a second attenuator of the MI², where the attenuation was adjusted by push-button switches controlled by the subjects. After the initial attenuation by the MI², the noise-bands were mixed in mixer 1, and then routed through an Hewlett-Packard attenuator (Model 350D). The noise-bands were then mixed with the shaped pure-tone in mixer 2. The composite stimulus was then low-passed filtered (Wavetek/Rockland; Model 752A) at 2300-Hz, amplified (SAE; A202), and passed through a final attenuator (Shallco; T-320-B), before being sent to a circumaural earphone (Sennheiser; Model HD430).
PART B (SEE FIGURE 19.)

The noise-bands and ramp were D/A converted by M208 modules of the M12. The noise-bands were then mixed in mixer 1, passed through the M108 attenuator module of the M12 (set to minimum attenuation), routed to an H-P 350D attenuator, and patched into mixer 2. At mixer 2, the noise-bands were mixed with the shaped pure-tone. Routing of the 140-ms ramp and pure-tone was identical to that described in part A of this experiment.

D. EXPERIMENTAL PROCEDURE

The levels of the noise-bands and the threshold procedure were the same in part A and part B of this experiment. Each noise-band in isolation was presented at 62 pSPL; except for those with envelope E, which were presented at 65 pSPL. Thresholds were determined using the tracking procedure described and used in experiment 1. Thresholds obtained in part A of this experiment were compared to those obtained in part B.

Although the same five pairs of noise-bands are used in both part A and part B, each pair in the strongly fused condition is considered separate from the same pair in the weakly fused condition. Therefore, ten different noise-band pairs are considered. In the analyses, the comparisons of particular interest are those between the noise-band pairs in the weakly fused condition and strongly fused condition comprising the same noise-bands.
FIGURE 19.

Instrumentation diagram for determining signal detection thresholds in correlated noise-band pairs, with the flanker-band and masker-band vertically-fused.
PART A: WEAKLY FUSED CONDITION.

As in experiment 2, a correlated noise-band was presented alternately with a noise-band presented in isolation. The noise-band presented in isolation was identical to the flanker-band of the noise-band pairs. The ISI was set at 15-ms. At this short ISI, the flanker-band should be captured into a horizontal-stream with the noise-band presented in isolation -- according to the findings of experiment 2. Capturing the flanker-band into a horizontal-stream should weaken vertical-fusion between masker-band and flanker-band.

There were 334 intervals in each trial. Therefore, in each trial, there were 167 occurrences of the pair of noise-bands, and 167 occurrences of the noise-band presented alone. The tone occurred with the co-occurring noise-bands only.

The duration of each trial was about 1 minute and 52 seconds. Ten threshold measures (replicates) were obtained in each of the five correlated noise-band pairs. See figure 20.

PART B: STRONGLY FUSED CONDITION.

Correlated noise-bands were presented at an ISI of 350-ms. A noise-band was not presented in isolation; therefore, fusion strength of the correlated noise-bands was not impaired. At this ISI, the time between the occurrences of
the pairs of noise-bands coupled with the pure-tone was identical to that which occurred in part A. Each trial consisted of 167 intervals. Therefore, the noise-band pair and pure-tone signal occurred the same number of times as in part A. Also, the duration of a trial (1 minute and 52 seconds) was the same as in part A. Ten replicates were obtained in each of the five correlated noise-band pairs. See figure 21.
FIGURE 20.
An illustration of the stimulus sequence for part A of experiment 3. The target-band (flanker-band) should be captured into a horizontal-stream with the captor-band.
An illustration of the stimulus sequence for part B of experiment 3. Notice that there is no captor-band; therefore, fusion strength should be high for the correlated noise-band pair.
RESULTS

A significant difference for threshold was found between the ten pairs of noise-bands (based on a 3 x 10 x 10, within-subject, repeated measures ANOVA, cond \[ F(9,18) = 5.81 \] p<.01; Appendix E). A level of statistical significance was not found across replications (reps \[ F(9,18) = 1.02 \] NS), nor was a significant interaction noted between replicates and noise-band pairs (cond * reps \[ F(81,162) = .93 \] NS); suggesting the absence of a learning effect. However, thresholds in the strongly fused condition did not differ significantly from thresholds in the weakly fused condition ( \[ F(1,18) = .56 \] NS; see contrast statement in appendix E and figure 22). In addition, none of the comparisons of particular interest (e.g., HH in strongly fused condition compared to HH in weakly fused condition) differed at a level of significance, according to Duncan's multiple range test (protection level =.05; appendix F; see figure 23).

DISCUSSION

In this experiment, signal detection levels in the noise-band pairs of the strongly fused condition occurred in the same order as those found in the correlated condition of experiment one. In experiment one, the lower thresholds in the correlated condition were associated with the correlated/uncorrelated
DETECTION THRESHOLDS FOR THE PURE-TONE SIGNAL IN THE WEAKLY FUSED AND STRONGLY FUSED CONDITIONS

FIGURE 22

The bar labeled "SF" represents the mean threshold level for the pure-tone signal in the strongly fused condition. The bar labeled "WF" shows mean threshold in the weakly fused condition. Bars indicate standard deviation of the data within each condition and across subjects.
DETECTION THRESHOLD FOR THE PURE-TONE SIGNAL IN EACH NOISE-BAND PAIR

FIGURE 23

Condition and noise-band pairs are identified on the abscissa. Mean threshold for a noise-band pair within a condition is denoted by the ordinate. Bars indicate standard deviation of the data within each noise-band pair across subjects.
comparisons wherein the largest CMR effects were observed. No such pattern emerged from the data obtained in this experiment, however.

Given that there was no statistical difference between thresholds in the weakly fused and strongly fused conditions, the results of this experiment may be interpreted as suggesting that auditory-grouping plays little or no role in the CMR phenomenon. However, suppose perceptual organization of the noise-band pairs in the strongly fused condition and the weakly fused condition did not differ sufficiently for a threshold difference to occur between the two conditions. In experiment one, it was suggested that a CMR effect did not occur in the comparisons wherein the uncorrelated noise-band pair had a relatively high coefficient of correlation, because a useful contrast in perceptual organization did not exist between the correlated and uncorrelated pairs within the comparisons. The same argument is offered here to explain no significant difference between thresholds in the two conditions of this experiment. It was alluded to earlier, that the percepts of horizontal-streaming and vertical-fusion are not necessarily binary (Bregman and Pinker, 1978; Bregman et al., 1985). In this experiment, capturing of the flanker-band into a horizontal-stream may not have decreased the strength of vertical-fusion of the correlated pairs to a level that created a functional contrast between perceptual organization of the co-occurring noise-bands in the strongly fused and weakly fused conditions. If this is so, the pure-tone would not have been more difficult to detect in the
"weakly" fused condition than in the strongly fused condition, and thresholds for
the pure-tone signal would be equivalent across the two conditions.

A second alternative explanation of why thresholds did not differ
significantly between the two conditions of this experiment considers differences
in the auditory cues that may have been available to subjects in part A and part
B of the experiment. In part A (weakly fused condition), the rapidly presented
sequence (15-ms ISI) created a simple melody composed of two alternating
pitches. Two of the three subjects (JC and MC) that participated in this
experiment reported that a quality change in this melody could be detected as
attenuation of the tone was reduced. These two subjects were able to use a
perceived change in melody as a cue for tone presence. In part B (strongly
fused condition), two correlated noise-bands were presented in every interval at
a comparatively long (350-ms) ISI. The pair of noise-bands did not alternate
with a noise-band presented alone, as they did in the weakly fused condition.
Therefore, the simple two pitched melody perceived in part A of this experiment
was not perceived here, and detection of the tone could not be assisted by
melody change.

On the one hand, according to subject reports, subjects were provided
information (melody cue) to assist them in detecting the presence of the pure-
tone signal in the weakly fused condition, that was not provided to them in the
strongly fused condition. Despite this added cue, signal detection thresholds
were not significantly better in the weakly fused condition than in the strongly fused condition.

On the other hand, according to the auditory-grouping hypothesis of CMR, if the masker-band is strongly fused with the flanker-band, a tone embedded in the masker should be detected at a lower decibel level than if the masker-band and flanker-band are not strongly fused. In this experiment, however, thresholds were not significantly better in the strongly fused condition than in the weakly fused condition.

Perhaps tone detection was assisted in the weakly fused condition by a melody cue, and was assisted in the strongly fused condition by auditory-fusion of the correlated noise-bands. If so, and the amount of assistance offered by both "mechanisms" was nearly equivalent, then thresholds should not have differed across the two conditions. As stated, they did not.
The intent of the three experiments constituting this dissertation was to determine if the psychoacoustic phenomenon called comodulation masking release (CMR) results from a more general process of the auditory system to perceptually organize its acoustic environment. In short, the purpose of this work was to test the auditory-grouping hypothesis of CMR.

The auditory grouping hypothesis of CMR is based on the following four assumptions: 1) spectrally separated bands of noise are strongly fused by the auditory system if the temporal envelopes of these bands of noise fluctuate in synchrony; 2) if the temporal envelopes of spectrally separated bands of noise fluctuate entirely independent of each other, the noise-bands will not be fused by the auditory system; 3) perceptual isolation of a static pure-tone centered in an AM masker will increase if the AM masker is fused with noise occurring outside the spectral region (i.e., critical-band) of the pure-tone; and 4) the greater the perceptual isolation of the pure-tone signal from the masker, the lower the decibel level necessary for the tone to be detected.

The first two assumptions of the auditory-grouping hypothesis were best addressed by experiment two. In this experiment, vertical-fusion was placed into competition with horizontal-streaming. That is, the stimuli were presented in a manner such that one noise-band of a noise-band pair could remain vertically-fused with the co-occurring noise-band, or it could be captured into a
horizontal-stream. It was reasoned that the stronger the vertical-fusion of a co-
occurring pair, the more difficult it would be to capture one of the noise-bands
into an horizontal-stream. In experiment two, subjects shortened the duration of
the inter-stimulus-interval (ISI) between a pair of noise-bands and a noise-band
presented alone (captor-band) until they believed the predominant perception to
be one of horizontal-streaming. It was found that shorter ISIs were needed to
capture a noise-band into a horizontal-stream if the noise-band was part of a
correlated noise-band pair, than if the same noise-band was part of an
uncorrelated noise-band pair. (The noise-bands captured into a horizontal-
stream were centered at 1700-Hz, and served as flanker-bands in experiments
one and three.) From this finding, it was concluded that the correlated noise-
band pairs were more strongly fused than the uncorrelated noise-band pairs.
This suggests that strength of fusion of co-occurring noise-bands may play a
role in the CMR phenomenon, and, thus, supports the first two assumptions of
the auditory-grouping hypothesis of CMR.

The results of experiment one were consistent with those of experiment
two. That is, they too suggested that fusion strength of AM noise-bands may
underlie the level differences in threshold for a pure-tone in the correlated and
uncorrelated noise-band conditions of CMR experiments. Experiment one
revealed an overall CMR effect of just under 3 dB. However, the magnitude of
the CMR effect was found to differ substantially across the five specific
correlated/uncorrelated noise-band comparisons. The magnitude of the CMR
appeared closely related to the degree of correlation of the temporal envelopes of the "uncorrelated" noise-band pairs (see footnote 1). Uncorrelated pairs with correlation coefficients closest to zero were associated with the specific comparisons that produced the largest CMR effects. Those uncorrelated pairs with the higher absolute values for correlation were associated with the specific comparisons that produced the lowest or absent CMR effects. It was suggested that the auditory system may have fused (to a degree) the uncorrelated noise-band pairs with the higher absolute values for correlation. If so, a contrast in the perceptual organization of the correlated noise-band pairs and the "uncorrelated" noise-band pairs would not have occurred. Without contrast, or functional contrast, between perceptual organization of correlated and uncorrelated noise-band conditions, perceptual isolation of a pure-tone signal centered in one of the noise-bands would not differ across the two conditions either. It was therefore proposed that degree of correlation of "uncorrelated" noise-bands influenced the degree to which the auditory system fused the uncorrelated noise-bands. This supposition was supported in experiment two, where it was found that the "uncorrelated" noise-band pairs with the higher coefficients of correlation (absolute values) required the shorter ISIs for capturing one of the noise-bands from the pair into a horizontal-stream. This suggested that the uncorrelated noise-band pairs with the higher coefficients of correlation were fused more strongly than those uncorrelated pairs with the lower coefficients of correlation.
It was also found in experiment one, that the lower signal detection thresholds in the correlated condition were in the correlated pairs that were involved in the correlated/uncorrelated comparisons from which the largest CMR effects emerged. The temporal envelopes of the masker-bands in these comparisons had very large "dips" during the time interval the tone was present. This finding was consistent with the work of Grose and Hall (1989) that showed the tone must occur in the dips of the masker for a CMR to occur, and also consistent with the "dip-listening" hypothesis of CMR proposed by Buus (1985). Although this finding is consistent with the dip-listening hypothesis of CMR, it does not rule out an hypothesis of auditory-grouping contributing to the CMR effect.

In experiment three, an attempt was made to weaken the fusion strength of correlated noise-band pairs without altering the acoustics of the pair. This was done by capturing the flanker-band into an horizontal-stream. Detection thresholds for the tone were obtained in the "weakly fused" condition, and then compared to thresholds in a "strongly fused" condition. This was done to determine if changes in fusion strength, alone, of correlated noise-bands influenced detection of the pure-tone signal. According to the assumptions of the auditory-grouping hypothesis of CMR, the signal should have been more difficult to detect in the weakly fused condition than in the strongly fused condition. However, there was no statistical difference for threshold
between the two conditions. This finding did not support the auditory-grouping hypothesis.

Two alternative explanations addressing lack of a threshold difference between the weakly fused and strongly fused condition were offered. The first suggests that thresholds differences did not occur because the fusion strength was not weakened sufficiently to create functional contrast in perceptual organization of the weakly fused and strongly fused condition. This same argument was offered to explain the lack of a CMR effect in the comparisons of experiment one where the coefficient of correlation for the uncorrelated noise-bands was farther from zero than in those comparisons where large CMR effects were observed. The second alternative explanation suggest that thresholds did not differ between the weakly fused and strongly fused conditions because an auditory cue based on melody change was available in the weakly fused condition that was not available in the strongly fused condition. It was suggested that this additional melody cue may have counter-balanced any effects on tone detection resulting from reduced fusion strength, causing thresholds in the weakly fused condition to be equivalent to those in the strongly fuse condition.

In summary, findings from the second experiment suggest that fusion strength is greater for noise-bands with correlated temporal envelopes than it is for noise-bands with uncorrelated temporal envelopes; the basic premise of the auditory-grouping hypothesis of CMR. The data from experiment one suggests
that the CMR effect does not occur if the uncorrelated noise-bands are fused by the auditory system. This is also in agreement with the auditory-grouping hypothesis. Therefore, although the results of experiment three did not support an hypothesis of CMR based on auditory-grouping, it appears too early to abandon this hypothesis.
BIBLIOGRAPHY


APPENDICIES
APPENDIX A

Summary of audiometric data for all subjects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>250</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
<th>6000</th>
<th>8000 (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JC</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>MC</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WJ</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PR</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>25</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>LM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Above values are thresholds, expressed in dBHL (Hearing Level), for air-conducted pure-tones presented through TDH-49 earphones.

<table>
<thead>
<tr>
<th>Subject</th>
<th>WRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>JC</td>
<td>100%</td>
</tr>
<tr>
<td>MC</td>
<td>100%</td>
</tr>
<tr>
<td>WJ</td>
<td>100%</td>
</tr>
<tr>
<td>PR</td>
<td>96%</td>
</tr>
<tr>
<td>LM</td>
<td>100%</td>
</tr>
</tbody>
</table>

Above percentages are word recognition performance scores for PB-50 wordlists. The monosyllabic words were presented to the test ear at 30 dBSL (Sensation Level), relative to the pure-tone-averages.
## APPENDIX B

ANOVA Summary Table for experiment one.

<table>
<thead>
<tr>
<th>Factor Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditions</td>
<td>9</td>
<td>1695.364</td>
<td>188.374</td>
<td>5.54*</td>
</tr>
<tr>
<td>Sub * Cond</td>
<td>18</td>
<td>612.502</td>
<td>34.028</td>
<td></td>
</tr>
<tr>
<td>Reps</td>
<td>4</td>
<td>22.334</td>
<td>5.584</td>
<td>.56</td>
</tr>
<tr>
<td>Sub * Reps</td>
<td>8</td>
<td>80.364</td>
<td>10.045</td>
<td></td>
</tr>
<tr>
<td>Condition * Reps</td>
<td>36</td>
<td>304.255</td>
<td>8.452</td>
<td>.91</td>
</tr>
<tr>
<td>Subject * Cond * Reps</td>
<td>72</td>
<td>667.599</td>
<td>9.27</td>
<td></td>
</tr>
<tr>
<td>(Contrast)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COR vs UNCOR</td>
<td>1</td>
<td>287.32</td>
<td>287.32</td>
<td>30.99*</td>
</tr>
<tr>
<td>Subject * Cond * Reps</td>
<td>72</td>
<td>667.599</td>
<td>9.27</td>
<td></td>
</tr>
</tbody>
</table>

* Significant beyond the .01 level
APPENDIX C

Duncan’s Multiple Range Test for Experiment One.

**Duncan’s Multiple Range Test for variable: DATA**

*NOTE:* This test controls the type I comparisonwise error rate, not the experimentwise error rate

\[ \text{Alpha} = 0.05 \quad df = 18 \quad \text{MSE} = 34.02787 \]

Number of Means 2 3 4 5 6 7 8 9 10


Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Duncan Grouping</th>
<th>Mean</th>
<th>N</th>
<th>COND</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>48.133</td>
<td>15</td>
<td>U E</td>
</tr>
<tr>
<td>A</td>
<td>46.287</td>
<td>15</td>
<td>C E</td>
</tr>
<tr>
<td>A</td>
<td>45.633</td>
<td>15</td>
<td>C A</td>
</tr>
<tr>
<td>B</td>
<td>44.627</td>
<td>15</td>
<td>U D</td>
</tr>
<tr>
<td>B</td>
<td>42.700</td>
<td>15</td>
<td>C H</td>
</tr>
<tr>
<td>B</td>
<td>41.560</td>
<td>15</td>
<td>C C</td>
</tr>
<tr>
<td>B</td>
<td>41.240</td>
<td>15</td>
<td>U H</td>
</tr>
<tr>
<td>F</td>
<td>40.253</td>
<td>15</td>
<td>U C</td>
</tr>
<tr>
<td>F</td>
<td>37.833</td>
<td>15</td>
<td>C D</td>
</tr>
<tr>
<td>F</td>
<td>37.667</td>
<td>15</td>
<td>C A</td>
</tr>
</tbody>
</table>

Taken from SAS printout.
APPENDIX D

ANOVA Summary Table for Experiment Two.

<table>
<thead>
<tr>
<th>Factor Variation</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditions</td>
<td>9</td>
<td>138530.322</td>
<td>15392.258</td>
<td>3.84*</td>
</tr>
<tr>
<td>Sub * Cond (ERROR)</td>
<td>36</td>
<td>144283.815</td>
<td>4007.884</td>
<td></td>
</tr>
<tr>
<td>Reps</td>
<td>7</td>
<td>11283.038</td>
<td>1611.863</td>
<td>.19</td>
</tr>
<tr>
<td>Sub * Reps (ERROR)</td>
<td>28</td>
<td>240058.425</td>
<td>8573.515</td>
<td></td>
</tr>
<tr>
<td>Condition * Reps</td>
<td>63</td>
<td>77033.938</td>
<td>1222.761</td>
<td>.84</td>
</tr>
<tr>
<td>Subject * Cond * Reps (ERROR)</td>
<td>252</td>
<td>368003.225</td>
<td>1460.330</td>
<td></td>
</tr>
<tr>
<td>(Contrast) COR vs UNCOR 1</td>
<td></td>
<td>87054.503</td>
<td>87054.503</td>
<td>59.61*</td>
</tr>
<tr>
<td>Subject * Cond * Reps (ERROR)</td>
<td>252</td>
<td>368003.225</td>
<td>1460.330</td>
<td></td>
</tr>
</tbody>
</table>

Significant beyond the point .01 level.
APPENDIX E

ANOVA Summary Table for experiment three.

<table>
<thead>
<tr>
<th>Factor Variation</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditions</td>
<td>9</td>
<td>1986.738</td>
<td>220.738</td>
<td>5.81*</td>
</tr>
<tr>
<td>Sub * Cond</td>
<td>18</td>
<td>684.408</td>
<td>38.023</td>
<td></td>
</tr>
<tr>
<td>(ERROR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reps</td>
<td>9</td>
<td>87.854</td>
<td>9.762</td>
<td>1.12</td>
</tr>
<tr>
<td>Sub * Reps</td>
<td>18</td>
<td>172.664</td>
<td>9.59</td>
<td></td>
</tr>
<tr>
<td>(ERROR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition *</td>
<td>81</td>
<td>659.003</td>
<td>8.13</td>
<td>.93</td>
</tr>
<tr>
<td>Reps</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond * Reps</td>
<td>162</td>
<td>1415.074</td>
<td>8.73</td>
<td></td>
</tr>
<tr>
<td>(ERROR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Contrast)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FUSED VS STREAMED</td>
<td>1</td>
<td>21.174</td>
<td>21.174</td>
<td>2.42</td>
</tr>
<tr>
<td>Subject *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond * Reps</td>
<td>162</td>
<td>1415.074</td>
<td>8.73</td>
<td></td>
</tr>
<tr>
<td>(ERROR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant beyond the .01 level.
VITAE

Patrick W. Rappold

Address: 4546 Alvin Dark, Apt. 28
Baton Rouge, LA 70802
Home phone: (504) 769-2193
Work phone: (504) 388-2545
Date of birth: March 29, 1960

Education:

1991 Louisiana State University, Baton Rouge:
Doctoral Candidate in Hearing Science/Audiology.
Secondary area: Business management
Dissertation: An investigation of the contributions of auditory-grouping to the comodulation masking release effect.

1985 Louisiana State University Medical Center, New Orleans:
Masters of Communication Disorders (M.C.D.).
Audiology.

1983 Louisiana State University, Baton Rouge:
B.S.
Communication Disorders.

Research Experience:

1987-present Research Assistant to John K. Cullen, Jr., Ph.D., and M. Jane Collins, Ph.D. Involved in research projects investigating the affects of perceived motion on sound lateralization, and the auditory processing of high-frequency signals.

1985-1986 Research Assistant to Kresge Hearing Research Laboratory of the South. Involved in projects investigating ABR during optokinetic stimulation, use of insert earphones for ABR testing, and speech discrimination of cochlear implant recipients.
Candidate: Patrick W. Rappold

Major Field: Communication Disorders

Title of Dissertation: An investigation of the Contribution of Auditory - Grouping to the comodulation masking release effect.

Approved:

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

M. Jan Clee

Ray Danleff

Jack S. Daniel

Paul A. Hoffman

Date of Examination:

April 22, 1991