Perceptual Effects of Auditory Distortion Products on Three-Tone Stimuli

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

in Psychology

by

Ewa Maria Borucki

Dissertation Committee:
Associate Professor Bruce Berg, Chair
Professor Virginia Richards
Professor Kourosh Saberi

2014
DEDICATION

To

Benjamin Yessayian for his unwavering love, support and encouragement over our 15 years together. It means everything.

And to my parents who taught me that a life spent learning is a life well lived-- thank you for supporting me, unconditionally, in every way.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT OF THE DISSERTATION</td>
<td>viii</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Distortion Products</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Background AM/QFM</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Empirical evidence for CDT effects in AM/QFM discrimination</td>
<td>5</td>
</tr>
<tr>
<td>1.3.1 Stimuli</td>
<td>6</td>
</tr>
<tr>
<td>1.3.2 Modifications of the adaptive procedure</td>
<td>6</td>
</tr>
<tr>
<td>1.3.3 Results</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Rotating quadrant model (RQM): the basics</td>
<td>11</td>
</tr>
<tr>
<td>1.5 Case studies: the effects of phase &amp; frequency randomization</td>
<td>14</td>
</tr>
<tr>
<td>1.5.1 Stimuli</td>
<td>17</td>
</tr>
<tr>
<td>1.5.2 Results</td>
<td>17</td>
</tr>
<tr>
<td>2 Effects of duration on AM/QFM discrimination</td>
<td>23</td>
</tr>
<tr>
<td>2.1 Subjects</td>
<td>23</td>
</tr>
<tr>
<td>2.2 Stimuli</td>
<td>24</td>
</tr>
<tr>
<td>2.3 General Procedure</td>
<td>24</td>
</tr>
<tr>
<td>2.4 Results</td>
<td>25</td>
</tr>
<tr>
<td>3 Monotic vs diotic discrimination ability</td>
<td>30</td>
</tr>
<tr>
<td>3.1 Subjects</td>
<td>30</td>
</tr>
<tr>
<td>3.2 Procedures</td>
<td>30</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>31</td>
</tr>
<tr>
<td>4 Effects of amplitude on the phase of the CDT</td>
<td>35</td>
</tr>
<tr>
<td>4.1 Predictions</td>
<td>35</td>
</tr>
<tr>
<td>4.2 Subjects</td>
<td>36</td>
</tr>
<tr>
<td>4.3 Stimuli</td>
<td>36</td>
</tr>
<tr>
<td>4.4 Design &amp; Procedure</td>
<td>36</td>
</tr>
<tr>
<td>4.5 Results</td>
<td>37</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5 Testing the RQM and moving beyond AM and QFM</td>
<td>40</td>
</tr>
<tr>
<td>5.1 Subjects</td>
<td>41</td>
</tr>
<tr>
<td>5.2 Stimuli</td>
<td>41</td>
</tr>
<tr>
<td>5.3 Design &amp; Procedure</td>
<td>41</td>
</tr>
<tr>
<td>5.4 Results</td>
<td>42</td>
</tr>
<tr>
<td>6 Discussion</td>
<td>44</td>
</tr>
<tr>
<td>References</td>
<td>48</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Simulation of random responses</td>
<td>8</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>AM/QFM threshold functions</td>
<td>10</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Rotating quadrant model</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Simulations for temporal spectral cues</td>
<td>16</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Subject 1, randomization thresholds</td>
<td>18</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Subject 5, randomization thresholds</td>
<td>19</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Subject 6, randomization thresholds</td>
<td>21</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>2 groups of listeners for AM/QFM</td>
<td>26</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Effect of reducing duration</td>
<td>28</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Monotic vs diotic thresholds</td>
<td>32</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Effect of reducing carrier intensity</td>
<td>37</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Thresholds as a function of phase of the lower component</td>
<td>42</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my committee chair, Associate Professor Bruce Berg, who always inspires me to be curious and seek answers. Thank you for your guidance. I would also like to thank my committee members, Professor Virginia M. Richards and Professor Kourosh Saberi for their support over the years. I am grateful for the helpful suggestions and resources provided that helped make this possible.

I would like to thank Matthew Turner for his excellent advice during my first few years and Hisaaki Tabuchi for being an excellent collaborator and your contributions to Chapter 1 of this work. Chapter 1 (sections 1.3 – 1.5) of this dissertation is a reprint of the material that appears in Hearing Research. This research was supervised by Dr. Bruce Berg.

This work was supported by grant number NSF BCS-07464003.
CURRICULUM VITAE

Ewa Maria Borucki

Education

Ph.D. University of California, Irvine
Cognitive Science, 2014

BA University of California, Irvine
Psychology, 2007

Publications:


Borucki, E, and Berg, B.G. Dissociating spectral and temporal influences in an amplitude modulation and quasi frequency modulation discrimination task (In preparation; to be submitted to Hearing Research)

Poster Presentations:


Association Memberships

Student member of the Acoustical Society of America
Through a series of experiments, the possible influence of cubic difference tones (CDTs) was investigated in a task traditionally used to investigate the bandwidths of phase sensitivity. In each procedure, estimates of modulation depth necessary to discriminate amplitude modulated (AM) and quasi-frequency modulated (QFM) tones were measured. Threshold functions for the listeners were often non-monotonic, with sharp non-monotonicities observed at higher bandwidths (e.g. above 600 Hz for a 2000 Hz carrier). This is likely due to the generation of a CDT at the site of the lower sideband creating a salient spectral cue. The cue is shown to be degraded by randomizing phase of the carrier frequency (Tabuchi, Borucki, and Berg, 2012). However, when duration is limited, temporal discrimination ability is degraded while a smaller effect is found at the sites of probable distortion product effects. When the intensity of the center frequency is decreased from 80 dB to 40 dB SPL a greater number of non-monotonicities are observed for high modulation frequency conditions, reflecting changes in phase of the CDT that occurs with changes in intensity. Finally, thresholds
were measured as a function of the phase of the lower side band; thresholds support a
model that suggests listeners discriminate between tone complexes based on the
interaction of the CDT with the lower sideband at higher bandwidths.
Chapter 1

Introduction

1.1 Distortion products

Distortion products (DPs) are components that are not present in an exogenous stimulus. They can be internally generated by using a pair of tones, $f_1$ and $f_2$, where the frequency of $f_1 < f_2$. Psychophysical investigations of DP effects began with Zwicker (1955) and continued throughout the 1970s, predating Kemp's (1978) discovery of otoacoustic emissions. The primary empirical method of that period was the probe-tone, beat-cancellation technique in which a probe-tone was added a few Hz away from the expected frequency of a distortion product in order to create the perception of beats (e.g. Zwicker, 1955; Goldstein 1967). Listeners adjusted the amplitude and phase of a cancellation tone with the same frequency as the distortion product in order to cancel the perception of beats. The DP was assumed to have the same amplitude as the cancelation tone and a phase difference of $\pi$ radians. Many of the basic discoveries of this period, such as the dominance of the difference tone (DT) and particularly the cubic distortion tone (CDT), were later confirmed with audiometric measurements of
distortion product otoacoustic emissions (e.g. Harris et al., 1989; Gaskill and Brown 1990). Hall (1972) found that the relative amplitude of the cancellation tone decreased when $f_2/f_1 > 1.2$; this is in direct agreement with later distortion product otoacoustic emission (DPOAE) measurements which found that the relative amplitude of emissions are greatest for a $f_2/f_1$ ratio of ~1.22. 1.22 is an estimate, the exact ratio varies by person and typically falls between 1.15 and 1.25 (Harris et al., 1989; Gaskill and Brown 1990).

In addition to the level of CDT being dependent on frequency separation of $f_1$ and $f_2$, the level of CDT is also affected by $L_2$. Goldstein (1967b) showed that the cancellation tone nearly proportionally increases as a function of $L_2$ up to around 30 to 40 dB SL at a for $L_2$ fixed at 50 dB SL. Smoorenberg (1972b) systematically examined the cancellation tone as a function of $L_2$ and $L_2$. It almost proportionally increases as a function of $L_2$ when $L_2$ is fixed. For example, the cancellation tone increases up to 30 dB when $L_2$ increases to 40 dB at a fixed $L_2=52$ dB which are consistent with Goldstein’s (1967b) results. It has also found that the CDT can be detected when primaries are only 20 dB above threshold. The CDT tends to be within 20 dB of the primaries when $f_2/f_1=1.2$ (Hall, 1972a; Smoorenburg, 1972a)

Since that time, in the psychophysical domain, distortion product effects have often been viewed as something to avoid or mask. The presence of these combination tones poses a unique problem for psychophysical experiments using tones by creating an uncontrolled confound. For instance, the discrimination of amplitude modulated
tones (AM) and quasi-frequency modulated tones (QFM), an ostensibly pristine technique for estimating phase sensitivity, is thought to be “contaminated” by distortion products (Goldstein 1967a; Buunen 1975).

1.2 Background AM/QFM

Mathes and Miller (1947) were among the first to investigate phase sensitivity using AM and QFM tones. In theory, an AM/QFM discrimination task is ideal for evaluating phase sensitivity as each stimulus has the same power spectra composed of three tones but having different envelopes as a result of a relative phase difference of the center tone. The center tone (i.e. carrier frequency) of the QFM stimulus is presented $\pi/2$ radians out of phase from the two sideband tones, whereas for AM, all three tones have the same phase. Consequently, QFM has an envelope modulation rate twice that of AM and a lower modulation depth. They found AM and QFM were found audibly different from each other when the frequency spacing between components was narrow, but became indistinguishable when the ratio of modulation rate to center frequency ratio was greater than 0.4. They proposed that auditory filters constitute the first stage of processing, an idea similar to Fletcher (1940), except that, rather than an integration of energy, “information” was transmitted by the temporal output of an auditory filter. According to their theory, AM and QFM can be discriminated only when the three tones are within the frequency range of an auditory filter. DPs, however, were suspected of contributing to the discrimination of AM/QFM tones In particular, the
interaction between the low frequency sideband and an internally generated CDT produced by the two higher tones were thought to generate a spectral cue (i.e. intensity of the lower sideband) that aided in discrimination (Buunen, 1975; Goldstein, 1967b). The CDTs generated by AM and QFM tones, designated $CDT_{am}$ and $CDT_{QFM}$, have different phases and thus interact differently with the lower sideband. Given that a CDT always interacts with the lower sideband and that it is the most audible distortion product (Plomp, 1965), a CDT effect is the most likely source of contamination due to potential distortion product effects.\(^1\)

To attempt to address this problem, Bernstein and Oxenham (2006) added lowpass noise in order to mask low frequency distortion products; however, no ideal cutoff frequency exists for such a technique because the CDT is always present at the frequency of the lowest tone of the stimulus. Without masking the CDT, there was no assurance that all pertinent distortion product effects were controlled; increasing the cutoff frequency of the noise in order to mask the CDT would have also masked the lower sideband of the stimulus affecting the temporal envelope.

Nelson (1994), on the other hand, claimed that auditory distortion products are too weak to produce an intensity cue in AM/QFM discrimination. Nelson measured percent correct for AM/QFM discrimination and found that at higher intensities maximal bandwidth for phase sensitivity is broader than for lower intensity stimuli.

\[1 \text{ The frequency of the CDT is defined as } 2f_1-f_2. \text{ In the case of AM and QFM stimuli, the frequency of the lower of the three sidebands can be expressed as } F_L=2f_c-f_l. \text{ Of the three tones that comprise either the AM or QFM stimulus, a CDT will always interact with } f_l \text{ due to the equal spacing between components. CDTs generated by other component combinations do not interact with AM or QFM stimuli.} \]
However, Nelson’s data had some irregularities. In some cases, at lower intensity levels, percent correct would improve from chance performance as bandwidths increased. Irregularities found in Nelson’s data tend to diminish at higher intensities, resulting in a higher percent correct and a wider estimated critical bandwidth for phase sensitivity. Nelson concluded that distortion products may induce envelope differences, but the magnitude of the CDTs themselves should not provide enough energy to induce a level cue; however, no thorough evaluation for that claim was given.

Strickland and Viemeister (1997) used an adaptive procedure to measure the minimum modulation depth needed to discriminate between AM and QFM tones for a 4000 Hz carrier, as a function of modulation rate, a procedure similar to those found in the following series of studies. Thresholds were flat until a bandwidth of 128 Hz. From this point, only two modulation rates are tested (256 and 512 Hz) Data collected in their experiment are limited and further detail would make the role of temporal peripheral filtering clearer. Unlike the data Nelson’s experiment they did not find performance improvements at higher modulation frequencies.

1.3 Empirical evidence for CDT effects in AM/QFM discrimination

Tabuchi, Borucki and Berg (2012) measured thresholds for modulation depth required to discriminate AM/QFM stimuli were estimated as a function of carrier frequency and overall stimulus bandwidth. Temporal modulation transfer functions
(TMTFs; Viemeister, 1979) were obtained for different carrier frequencies at a single intensity.

1.3.1 Stimuli

AM and QFM stimuli are represented by:

\[
y(t) = \sin(2\pi f_C t + \theta) + \frac{m}{2}\left[\sin(2\pi (f_c + f_m) t)\right] + \frac{m}{2}\left[\sin(2\pi (f_c - f_m) t)\right] \tag{1.1}
\]

The stimulus defined by Eq. 1.1 is composed of a carrier, \(f_C\), and two sidebands, \(f_i=f_c-f_m\) and \(f_H=f_c+f_m\), where \(f_m\) is the modulation frequency and overall stimulus bandwidth is \(2(f_m)\). The stimulus \(y(t)\) is defined as AM when \(\theta\) is 0 and QFM when \(\theta\) is \(\pi/2\). Modulation depth, \(m\), of the side-bands is a dependent variable varying between 0 and 1. The value of \(m\) is titrated with an adaptive procedure in order to estimate the threshold pertaining to the AM and QFM discrimination.

1.3.2 Modifications of the adaptive procedure

In order to avoid over modulation (i.e. \(m>1\)) an upper limit of \(m=1\) was imposed on the titration of modulation depth with an adaptive procedure (Levitt, 1971). Previously, Strickland and Viemeister (1997) circumvented this problem by terminating a block of trials whenever the adaptive schedule dictated an \(m>1\). This technique is, however, insensitive to above-chance performance at high modulation
rates. In our procedure, when an incorrect response was obtained when \( m=1 \), \( m \) remained at 1 until two correct responses were obtained.

A second modification consisted of a non-conventional up-down schedule, \( m \) is decreased by 2 dB following two consecutive correct responses and is decreased by 2 dB following each subsequent correct response. This schedule was reset and \( m \) increased by 2 dB after each incorrect response. This procedure yields a probability of correct response lower than 0.71 and an analytical solution to the expected probability is lacking. Empirically, the probability of a correct response was found to be approximately 0.64 when the truncation count was zero, as shown in Fig. 1.1. The intent was to lower expected thresholds in order to reduce the number of truncation counts.

The occurrence of truncations when \( m=1 \) violated assumptions required for averaging across reversal points in order to calculate a threshold. Instead, all stimulus values after the first four reversal points were averaged and taken as a threshold (Klein, 2001)\(^3\). The chance level of performance was estimated by simulations of a model of random responses. The probability distribution of 3000 simulated thresholds and the corresponding cumulative function are shown in Fig 1.1. Five percent of the thresholds are less than -8.62 dB (20logm).
1.3.3 Results

Thresholds \([20 \log(m)]\) are plotted as a function of stimulus bandwidth in Fig. 1.2, with each row representing the data for an individual and each column representing a different carrier frequency. The dotted line at -8.62 represents chance performance estimated from simulations of random responses. Thresholds tend to a relative minimum in the vicinity of 100-200 Hz. Some listeners have elevated thresholds at the lowest modulation rates likely due to the relatively brief stimulus duration of 200 ms (see Viemeister, 1979). Although notably there are large individual differences at the narrowest of bandwidths, these data for these narrow bandwidths are
of secondary interest. One listener (S2) shows reduced phase sensitivity for all
carriers, with chance performance for all observed bandwidths with a 500 Hz carrier.

Data are often nonmonotonic with non-monotonicities found at wide stimulus
bandwidths. The range of individual differences is considerable. With the exception of
S2, all listeners show non-monotonicities beyond a bandwidth of 400 Hz for at least one
carrier frequency. For $f_c = 2000$ Hz, S3, S1, and S5 exhibit what can be described as a
double-dip non-monotonicity and all subjects but S2 demonstrate some kind of non-
monotonicity. In a number of cases, thresholds at these higher bandwidths are
substantially lower than the plateau of the TMTF.
Figure 1.2: Thresholds for modulation depth plotted as a function of stimulus bandwidth. Individuals and carrier frequencies represented in rows and columns respectively. The dotted lines represent chance performance. The asterisk indicates the conditions repeated for case studies. Standard error is shown with error bars.
This preliminary experiment demonstrates the inherent problem of estimating phase sensitivity with an AM/QFM discrimination task using tones. While it is certain that discrimination ability is driven by the phase difference between carrier components, it is less clear what cue subjects are using to discriminate between AM and QFM. For example, the threshold for S3 with $f_c=500$ Hz and a bandwidth of 400 Hz almost certainly reflects the effects of distortion products and can be confidently disregarded as evidence of phase sensitivity at that bandwidth. For S5 the non-monotonicities of the TMTF at high modulation rates with $f_c=1000$ Hz and for S6 with $f_c=2000$ Hz are more questionable. Despite some points of ambiguity, the evidence for distortion product effects at high modulation rates is persuasive.

1.4 Rotating quadrant model (RQM): the basics

An attempt to describe the potential effects of CDTs on AM/QFM discrimination has been put forth by Tabuchi, Borucki and Berg (2012). According to the rotating quadrant model (RQM) the interaction between the CDT and lower side band of either the AM or QFM stimulus, $f_l$, introduces an intensity cue in AM and QFM discrimination. The model draws on three basic findings from probe-tone, beat-cancellation experiments. First, $CDT_{am}$ and $CDT_{qfm}$ should differ in phase only with the relative phase difference between $CDT_{am}$ and $CDT_{qfm}$ being $\pi/2$ radians. Second, the absolute phase of each increases as the frequency separation between the carrier ($f_c$) and upper side band ($f_u$) increases (Hall, 1972a,b), and most importantly, the CDT interacts with the lower
tone of the stimulus (Buunen, 1975). In terms of vector addition, it is useful to visualize two equal-length vectors representing the CDTs forming a quadrant positioned “on top” of a longer vector representing $f_L$ (see figure 3.1). The distance from each respective “tip” of the quadrant to the base of $f_L$ represents the combined amplitude of the CDT and lower sideband.

The CDT generated by $f_C$ and $f_H$ interacts with $f_L$, and the phases of $CDT_{am}$ and $CDT_{qfm}$ are $\pi/2$ radians out of phase, forming a right angle quadrant. The quadrant rotates as $f_H$ increases in frequency. In a non-detect state, $CDT_{am}$ and $CDT_{qfm}$ add to $f_L$ in such a way that $f_L + CDT_{am}$ and $f_L + CDT_{qfm}$ are identical in length and contribute to AM/QFM discrimination equally. As $f_H$ increases, $CDT_{am}$ and $CDT_{qfm}$ rotate synchronously as a function of $f_H$, preserving their inherent phase relationship. When these two vectors are rotated $f_L + CDT_{qfm}$ are no longer equal to $f_L + CDT_{am}$, leading to a detect state.

Without precise knowledge about the phase and intensity of CDTs for individual listeners, the RQM has little predictive power and provides only post hoc qualitative descriptions of threshold functions. The model is nevertheless useful in considering the range of individual differences. According to the RQM, the magnitude of a non-monotonicity depends on the amplitude of the CDT and the location of the minimum threshold depends on the phase of the CDT. Probe-tone cancellation experiments show

---

2 Note that the RQM does not estimate the amplitude of the CDT relative to the primaries. Also, the strength of a CDT is dependent on the frequency separation between tones, being most robust at approximately 1.22 fh/fc.
much across-listener variability in estimates of these parameters (Buunen, Festen, Bilsen, and van den Brink, 1974; Zwicker, 1981; Zwicker and Fastl, 1999), and so the observed range of individual differences is actually expected.

Figure 1.3: Rotating quadrant model. States (a) and (c) represent non-detect states as $CDT_{am} + f_L = CDT_{qfm} + f_L$. Phases of both $CDT_{am}$ and $CDT_{qfm}$ advance synchronously as a function of frequency separation between $F_C$ and $f_H$, whereas the relative phase difference between $CDT_{am}$ and $CDT_{qfm}$ is $\pi/2$. States (b) and (d) represent detect states. From Tabuchi, Borucki and Berg (2012).
1.5 Case studies: the effects of phase & frequency randomization

A more important use of the RQM is a quantitative test of assumptions underlying the technique for degrading the presumed intensity cue as shown here by Tabuchi, Borucki and Berg (2012). Simulations of the RQM show that trial-by-trial randomization of the phase of the carrier over a modest range (e.g. –π/3 to π/3) degrades intensity information, whereas the same range of phase randomization produces only a modest effect on information conveyed by the envelope of the waveform. The vector difference between $CDT_{am}$ and $CDT_{qfm}$ based on the RQM (a) and the simulated threshold from temporal decision variables (b) as a function of randomized phase range. Panel (a) plots the dB difference between the AM and QFM vectors. Circles and diamonds respectively represent cases in which the CDT is 20 dB and 15 dB below the level of $f_L$. Panel (b) shows simulated thresholds for the decision variable, modulation depth or modulation rate.

The level difference between AM and QFM can be expressed as

$20\log(f_L+CDT_{am}/f_L+CDT_{qfm})$, where $f_L+CDT=(f_L^2 + CDT^2 + 2f_L + CDTCos\theta)^{1/2}$ and $\theta$ is the angle between the two vectors. The maximum difference between $f_L + CDT_{am}$ and $f_L+CDT_{qfm}$ is found when the respective phases of the CDT vectors are $\pi/4$ and $3\pi/4$ radians (or $-3\pi/4$ and $-\pi/4$) radians. Using the assumption that the intensity of the CDT is 20 dB lower than the intensity of $f_H$ (Zwicker and Fastl, 1999), then the maximum level difference is 1.2 dB, a value slightly above a JND for a pure tone (Jesteadt et
Fig. 1.4 (a) shows the mean level difference across 5000 simulated trials as a function of the range of phase randomization. The mean remains above 1 dB for randomization ranges of up to $2\pi/3$ radians, but the substantial increase in the standard deviation implies the level cue becomes unreliable even for small perturbations in the carrier phase. Temporal information is more resilient to phase randomization. Fig. 1.4(b) shows thresholds for two simulated decision variables as a function of the range of phase randomization. One is the max/min ratio of the Hilbert envelope (Forrest and Green, 1987), representing a depth of modulation cue. The other compares two-point power spectra of the envelope at $f_m$ and $2f_m$, representing a modulation rate cue. For both the ‘depth’ and ‘rate’ decision variables, internal noise is added at the level of the waveform, $f(t) + e(t)$, where, $e(t)$ is a Gaussian variable. The variance of the internal noise is adjusted to yield performance in the constant-phase condition that is roughly equivalent to the thresholds of listeners and then fixed for simulations of the random phase conditions. Threshold functions for the two decision variables are essentially the same, demonstrating that the two decision variables cannot be distinguished with the current experiment. It is apparent that the information conveyed by the envelope is only slightly degraded by a moderate range of phase randomization. The intent is to degrade potential pitch cues (Buunen, 1975), but the level cue may also be affected by frequency randomization. If the phase of the CDT changes when the “primaries”, $f_H$ and $f_C$, are shifted equally in frequency, then the corresponding rotation of the RQM would degrade intensity information in the same manner as phase randomization.
Figure 1.4: Panel (a) shows simulated dB difference between the AM and QFM vectors for a with the level of the CDT either 15 or 20 dB below the level of \( f_i \) shown in circles and diamonds respectively. Panel (b) shows simulated thresholds for the decision variable, modulation depth or modulation rate. Replotted from Tabuchi, Borucki & Berg (2012).
1.5.1 Stimuli

Stimuli are the same as formula 1.1 with the exception of adding two random variables, $f_k$ and $\theta_k$.

$$y(t) = \sin\{2\pi(f_C+f_k)t+(\theta+\theta_k)\} + \frac{m}{2} \sin\{2\pi(f_{C+m}+f_k)t\}$$

$$+ \frac{m}{2} \sin\{2\pi(f_{C-m}+f_k)t\}.$$  

(1.2)

The frequency shift, $f_k$, was sampled from uniform distribution with a range of -10 to 10 ($\Delta f=20$ Hz) or with a range of -40 to 40 ($\Delta f=80$ Hz). Likewise, the phase shift, $\theta_k$, was sampled from uniform distribution with a range of $-\pi/12$ to $\pi/12$ ($\Delta \theta=\pi/6$) radians, or with a range of $-\pi/6$ to $\pi/6$ ($\Delta \theta=\pi/3$) radians. The two random variables, $f_k$ and $\theta_k$, were independently sampled for AM and QFM on every trial. All other aspects such as stimulus generation and procedures were identical to the first experiment.

1.5.2 Results

In the previous experiment shown in Chapter 1.3, the lowest thresholds are found for conditions that also have non-monotonic qualities at wide bandwidths. Data from the 2000 Hz condition of the first experiment are reproduced with dotted lines in Fig.1.5 for S1. Data from randomization conditions are represented by the solid lines.

Without randomization, thresholds for S1 yield minima at bandwidths of 670 Hz and 940 Hz performance at intermediate bandwidths for both listeners. As mentioned previously, this pattern is consistent with one full-rotation of the RQM (see Fig. 1.3).
Thresholds at the minimum of the first non-monotonicity are lower than thresholds at the minimum of the second non-monotonicity.

Figure 1.5: Solid lines indicate functions for thresholds with randomization. Data from experiment one, with no randomization, are shown using a dotted line. The dotted line at -8.62 indicates chance performance. Carrier frequency and randomization parameters are shown at the top of each panel. Plot originally from Tabuchi, Borucki and Berg (2012).
The solid lines in the upper panels of Fig. 1.5 show thresholds with $\Delta f=20$ Hz and $\Delta \theta = \pi/6$. The minimum threshold of the first non-monotonicity is unaffected by the randomization, but is more sharply “tuned” in the sense that thresholds for points adjacent to the minimum are now at chance performance. The second non-monotonicity is no longer present and thresholds for bandwidths less than 400 Hz show a maximum increase of five decibels, with many points unaffected. The bottom panel of Fig. 1.5 shows the effects of increasing the phase randomization to $\Delta \theta= \pi/3$ with no frequency randomization. This level of randomization eliminates the non-monotonicities for S1.

Figure 1.6: Case study of S5. See caption for Fig. 1.5.
S5 exhibits distinctive non-monotonicities, with non-monotonicities spanning a relatively broad region of bandwidths for \( f_c = 2000 \) Hz and an extremely low threshold for \( f_c = 4000 \) Hz, as shown by the dotted lines in Fig. 1.6. S5 also gives an anomalous result in which performance improves with randomization is also found. For \( f_c = 2000 \) Hz and a randomization level of \( \Delta \theta = \pi / 6 \) and \( \Delta f = 20 \) Hz (upper-left panel), the threshold at 800 Hz decreases by 5 dB in comparison to the baseline condition. The low threshold at an adjacent point, 670 Hz, helps establishes the consistency of this finding. There also appears to be a narrowing of the non-monotonicity which indicates that the randomization is affecting a CDT-induced cue suggesting that the effect is not attributable to practice effects. A possible explanation is suggested by data from Hall’s (1972b) two-tone, probe cancellation experiment. Generally, the phase of the CDT is a linear function of the phase of \( f_H \), a foundational assumption of the RQM, but in a few instances reported by Hall, the function relating the phase of the CDT to \( f_H \) inexplicably resembles a step function. One implication is that points near the edge of the step are unstable so that a small, random change in the phase of \( f_H \) produces a large shift in the phase of the CDT which induces a detect state. This odd effect illustrates that the effects of nonlinearities may not yet be entirely predictable, and perhaps offers a prelude to the level of explanation that might be gained by combining the current method with a probe-tone, beat-cancellation experiment. Results for S5 with \( f_c = 4000 \) Hz are shown in the right panels of Fig. 1.6. For bandwidths greater than 400 Hz, thresholds generally increase by about 10 dB when the randomization level is \( \Delta f = 20 \) and \( \Delta \theta = \pi / 6 \), followed by another 5 to 10 dB increase when the randomization is increased to \( \Delta f = 80 \) and
$\Delta \theta = \pi/3$. These findings are consistent with predictions of the RQM, demonstrating a cause-and-effect relationship between the amount of randomization and thresholds. For stimulus bandwidths less than 400 Hz, randomization has little effect.

Figure 1.7: Thresholds for S6; see caption to Fig 1.5.
Data for S6 is of interest because this listener generally exhibits the best sensitivity in the first experiment. TMTFs from the first experiment, reproduced with dotted lines in Fig. 1.7 for $f_c=2000$ Hz and $f_c=4000$ Hz, have an extended low-threshold plateau followed by a relatively shallow roll-off. Results from randomization conditions suggest that the shape of the function might be the product of an overlap between the bandwidth of phase sensitivity and the region of distortion product effects. The solid line in the upper panel of Fig. 1.7 shows the results for $f_c=2000$ Hz with $\Delta \theta=\pi/3$. The minimum threshold at a bandwidth of 800 Hz increases to chance performance with randomization, but the effect at 560 Hz perhaps speaks more to the point. The 6 dB increase in the threshold at this bandwidth is greater than expected if sensitivity is strictly dependent on phase sensitivity and is less than expected if sensitivity is strictly dependent on a CDT induced intensity cue. It is plausible that both cues contribute to and that phase sensitivity remains when the intensity cue is degraded by randomization. A similar occurrence is found for $f_c=4000$ Hz at a bandwidth of 400 Hz, shown in the lower panel. A degraded intensity cue may also underlie the threshold increase at 200 Hz, particularly because the lack of an effect at narrower bandwidths implies that phase sensitivity for this listener is largely unaffected by this level of randomization.
Chapter 2

Experiment 1: Effects of duration on AM/QFM discrimination

This study contributes complementary evidence that non-monotonicities in TMTFs are attributable to CDT effects. Whereas in the previous chapters spectral cues were degraded while preserving temporal cues, the approach taken here is to impair temporal discrimination ability while preserving spectral cues by reducing the stimulus duration. Reducing duration should reduce a listeners’ ability to discriminate between AM and QFM stimuli temporally, as they are given fewer looks at the stimulus envelope (Viemeister & Wakefield, 1991). On the other hand, reducing duration should have less of an effect on a spectral intensity cue (Watson and Gengel, 1969) such that that the non-monotonicities in TMTF’s are more resilient to the stimulus manipulation than thresholds at low modulation frequencies.

2.1 Subjects

Seven subjects participated in this experiment; all were affiliates of the University of California Irvine between the ages of 20 and 35. However, only 4 subjects,
those showing non-monotonicities as seen in the previous chapter, participated in the second part of the experiment. Listeners received monetary compensation for their participation. All listeners had normal pure-tone absolute thresholds within the range of stimulus frequencies used.

2.2 Stimuli

Stimuli were the same as given in formula 1.1.

2.3 General procedure

Given the overall reliability of finding a non-linearity for a center frequency of 2000 Hz (Tabuchi, Borucki and Berg, 2012) center frequency was fixed at 2000 Hz while the modulation frequency, $f_m$, was varied between 25 and 800 Hz. Subjects typically completed 5 to 10 blocks of 50 trials and 4-5 conditions during a 2 hour session, with conditions presented to subjects in random order in sets of 5 blocks. For all subjects, thresholds for modulation depth necessary to discriminate between AM and QFM stimuli were measured as a function of modulation rate. Subjects were trained until they produced stable thresholds (generally 2-3 hours). As in the previous chapters, thresholds for modulation depth, $m$, are measured with a modified procedure where $m$ is allowed to vary between 0 and 1. However, $m$ varies in a traditional one-up two-down adaptive procedure yielding 71 percent correct (Levitt, 1971), in 2 dB steps,
in order to estimate thresholds for modulation depth required to discriminate between AM and QFM. AM and QFM were presented in random order in a 2IFC with each order equally likely. Subjects indicated which interval they believed contained the AM stimulus by pressing 1 or 2 on a keyboard, and subjects were given accuracy feedback after every trial. Each trial consisted of two 200 ms presentation intervals with an inter-stimulus interval of 500 ms. Subjects were also shown a plot of the adaptive track after each block. Stimuli were gated by a 20 ms-squared cosine ramp at onset and offset. The level of the carrier was calibrated at 70 dB SPL. Sounds were generated digitally, converted to analog with an external converter (E-MU 0202 Audio/MIDI interface) at a sampling rate of 22.05 kHz, and presented diotically over headphones (Sennheiser eh350) to listeners seated in a single-wall sound attenuating chamber.

For four subjects, those with nonmonotonic functions, thresholds for AM/QFM discrimination subsequently were estimated for shortened duration stimulus presentations; duration was limited to 20 ms in most cases (for one subject a 50 ms condition was used). For these shorter durations, stimuli were gated by 5 ms.

2.4 Results

Threshold functions for each listener are shown for the 200 ms condition in Figure 2.1. Listeners are divided into two groups, those who exhibit a monotonic TMTF (upper panel) and those with an essentially non-monotonic threshold function (lower panel). For the most part, the monotonic TMTFs are similar to those obtained for
AM/QFM noise discrimination (e.g. Strickland & Viemeister, 2000; Strickland, 2000; Eddins 1999), showing lessening sensitivity as the separation between tones becomes larger. Thresholds increase until reaching a modulation frequency of approximately 300 Hz. The dotted line at -2.65 represents chance performance non-monotonicities were not as reliable in this group as found previously, however, four of seven do exhibit non-monotonicities. Also, there are far fewer individual differences in the monotonic group as the non-monotonic group.

Figure 2.1: TMTF for modulation depth as a function of modulation frequency. Subjects 1-3 are shown in the top panel with circles, triangles and diamonds respectively. Subjects 4-7 are shown in the bottom panel with open circles, diamonds, upside down and right-side up triangles respectively. The dashed line at -2.65 represents chance performance.
Even though not all subjects show non-linearities, some generalization can be made. The non-monotonic TMTFs in the lower panel show a similar decrease in sensitivity at for increasing modulation rates, after thresholds decrease to near-chance levels nearing 300 Hz, thresholds drop and reach a minimum at a bandwidth of ~800 Hz. The data show only modest deviation between subjects for the transition point. For example, for S5 this transition occurs at 250 Hz.

For the listeners exhibiting a non-monotonic TMTF, a local minimum is found at a modulation frequency of 400 Hz, with the exception of S7 whose minimum lies at 500 Hz. A modulation rate of 400 Hz corresponds to an $f_2/f_1$ ratio of 1.2 (i.e. $f_1 = 2000$ Hz and $f_2 = 2400$ Hz, where $f_1$ and $f_2$ represent $f_C$ and $f_H$ respectively).

As shown in Fig. 2.2, decreasing the stimulus duration has the greatest effects at low modulation rates. Open circles represent the thresholds obtained with a stimulus duration of 20 ms, which is compared to thresholds for the 200 ms condition replotted from Fig. 2.1 as closed squares. For S4, data is also presented for a 50 ms duration, represented by the open triangles. Listener S4 presents the clearest demonstration that the cues for discrimination are different for low and high modulation rates. When the modulation rate is less than 200 Hz, thresholds increase systematically as the stimulus duration is decreased. The segments of the TMTFs for rates less than 200 Hz shift in parallel for the three different durations to a level of chance performance. In comparison, reducing the duration to 50 ms has only a small effect on the non-monotonicity. The location of the minimum threshold shifts from 400 to 450 Hz with little change in the overall span and depth of the non-monotonicity. When the duration
is further reduced to 20 ms, the span of the non-monotonicity decreases, but there is little effect on the minimum threshold. S5 withdrew from the study before completing the 20-ms condition, but the data are consistent with the results for S4. The single point for S5 at a modulation rate of 75 Hz is at chance performance, whereas the two points at high modulation rates show only a slight increase in thresholds, and it may be the case that the increase simply reflects a shift in the minimum, as observed for S4.

Figure 2.2: Thresholds for modulation depth plotted as a function modulation rate. 200 ms duration presentations are represented with closed squares; 50 ms presentations are represented with open triangles, and 20 ms stimulus presentation are shown in open circles. Center frequency for all conditions is 2000 Hz.

S6 and S7 exhibit measurable sensitivity at low modulation rates when the stimulus duration is reduced to 20 ms. For S6, the corresponding segment of the TMTF
shows a roughly parallel upward shift similar to that found for S4, but S7 presents a
different case in that the increases in thresholds at 50 Hz and 100 Hz are much greater
than the slight increase found at 200 Hz. The result suggests that a CDT-induced
spectral cue is still available to S7 at 200 Hz when the temporal information is degraded
by decreasing the stimulus duration. Conversely, S7 also exhibits the greatest effects of
reducing the stimulus duration at high modulation rates in absolute terms, but one
must consider the exceptionally low minimum threshold of -40 dB for a modulation rate
of 500 Hz, 13 dB lower than the threshold at 100 Hz. Both S7 and S6 continue to
produce a robust non-monotonicity for the 20 ms duration.
Chapter 3

Experiment 2: Monotic vs diotic discrimination ability

3.1 Subjects:

4 subjects participated in this experiment. Subjects 1 and 2 also participated in the duration study and showed no ability to discriminate between AM and QFM at higher bandwidths. S4 also participated in the previous experiment, but had non-monotonic data. Subject 3 was a new subject and was trained for 4 hours prior to data collection.

3.2 Procedures:

Procedures were generally the same as in previous experiment. Stimulus presentations were given in 200 ms intervals with an isi of 500 ms. Thresholds were first measured diotically, then monaurally for both the left and right ears, and finally listeners who previously had no non-monotonicities discriminated between diotic AM and QFM stimuli for a second time.
3.3 Results:

As shown in Fig. 3.1, first presentation showed no or extremely modest distortion product effect at higher modulation frequencies for S1, S2 and S3. However, monaurally, all subjects were able to discriminate between AM and QFM tones at some higher modulation frequency, with all subjects producing some kind of non-monotonicity at expected higher modulation frequencies as seen in the previous chapters.

Data for monaural conditions, shown with open symbols, reflect an increase in sensitivity for S1, S2 and S3. Both S1 and S2 show large decreases in thresholds for modulation frequencies above 400 Hz.

This effect was weakest for S3 whose maximum improvement was 10 dB for a modulation frequency of 550 Hz. As a further point of interest, the minima found at higher modulation frequencies are not necessarily the same. For example, S3 shows a minimum at a modulation frequency of 550 Hz and 400 Hz for the left and right ear respectively. Thresholds obtained from the right and left ear differ between each ear for all subjects.
Figure 3.1: Thresholds for each subject are plotted as a function of modulation frequency. Initial diotic presentation filled circles are replotted from experiment 3 (with the exception of S3). Thresholds for left and right ear stimulus presentations are shown with open circles and triangles, respectively. Filled triangles represent the second diotic stimulus presentation. The dashed line at -2.65 indicates chance performance.
The second run of the diotic condition yielded somewhat surprising data. Thresholds are similar to those of the monaural conditions. As seen for subject S2, for the range of modulation frequencies collected, diotic thresholds generally reflect the sensitivity of the more sensitive ear for any particular condition. Only for S2, for $f_m = 450$ Hz were any sizable improvements seen for diotic stimulus presentations (showing an improvement of approximately 10 dB from the left ear condition). S1 shows some smaller improvements from the best ear; at 450 Hz a 3.6 dB improvement is found for the right ear, and at 500 Hz a 2.5 dB improvement is found for the left ear. Overall, data suggest a special case of training effect. Given that subjects ran 5-10 blocks of 50 trials for every condition on the initial diotic AM/QFM discrimination task, had relatively small standard errors and thresholds do not improve until presented with monaural stimuli, it appears that being exposed to monaural stimulation induces an ability to attend to differences between AM and QFM tones presented diotically at higher modulation frequencies. It does not appear to be the case that the proximal stimulus is the same for the right and left ears; therefore, while the stimulus is technically being presented to the listener diotically, it is more useful to think of the stimuli as being attended to dichotically.

To further illustrate the complexities the data, monaural data were collected for listener S4 (shown on the bottom panel of Fig. 3.1) who demonstrated a sizable non-monotonicity. Thresholds are notably different between the right and left ears with the left and right ears. Threshold functions for the left and right ears show a minimum at 400 and 500 Hz respectively. For a separation of 500 Hz between primaries, the diotic
condition produces the lowest overall threshold. In the light of complexities demonstrated here with diotic stimulation, all subsequent data were collected monaurally.
Chapter 4

Experiment 3: Effects of amplitude on the phase of the CDT

4.1 Predictions:

In earlier cancellation tone experiments, Hall (1972) found that the rate of phase change for a cancellation tone was inversely proportional to amplitude of the primaries. In Hall’s experiment $f_1$ was fixed 1475 Hz while $f_2/f_1$ ranged from 1.1 to 1.5 with primaries set to either 78 or 58 dB SPL. Hall found that for an increasing $f_2/f_1$, the phase of the cancellation tone for the lower intensity primaries advanced at nearly twice the rate of the higher intensity primaries. However, data are rather limited as Hall was his only subject. If data gathered from Hall are universal, the RQM would predict that the number of non monotonicities found on our threshold functions would also be inversely proportional to stimulus intensity; the rate of phase change of the CDT would be higher at low intensities for increasing stimulus bandwidths and thresholds would cycle between detect states and non-detect states faster. As the frequency separation between $f_c$ and $f_{1H}$ (i.e. $f_1$ and $f_2$ in Hall’s experiment) increases, the phases for
both \( CD_{am} \) and \( CD_{qfm} \) would advance faster for lower intensities (i.e. a faster rotation of the RQM).

4.2 Subjects

3 subjects participated and received course credit or monetary compensation for their participation.

4.3 Stimuli

3 listeners discriminated between AM/QFM for a carrier calibrated to 80 dB SPL, of whom two subjects also discriminated for a 40 dB SPL carrier. The level was only reduced to 60 dB SPL for the third listener as thresholds never stabilized for that listener for the 40 dB SPL condition\(^3\). Center frequency was held constant at 2000 Hz.

7.4 Design & Procedure

Procedures are largely similar to the previous experiments. Rather than being seated in a single walled sound attenuating chamber, subjects completed the experiment in double-walled sound attenuating chamber. All other equipment were kept the same. Data for the 80 dB condition were collected first, followed by data for the reduced intensities. All data were collected monaurally from subjects’ right ears.

---

\(^3\) S3 was simply not very good at the 40 dB SPL condition. For this subject averaging across the last 4 reversal points in a 50 trail block was problematic. The subject reported ‘losing’ the signal midway through a track. This is supported by the subject’s adaptive tracks which, more or less, tended to be a U shape function. The matter was easily resolved by raising the stimulus level.
Figure 4.1: Thresholds are shown as a function of modulation rate. 80 dB thresholds are shown in open circles and 40 dB (S1 & S2) and 60 dB thresholds are shown with filled squares. Center frequency is fixed at 2000 Hz. The dotted line at -2.65 represents chance performance.
Data for the 80 dB condition (top panel of Fig. 4.1) are shown in open circles whereas data for the 40 dB condition are shown in filled squares. For modulation frequencies above 300 Hz, data for S1 show non-monotonicities. For the 80 dB carrier, this listener shows a very broad region of sensitivity at high modulation frequencies spanning from 450 Hz to 800 Hz. Between frequency separations of 480 and 650 Hz thresholds reflect a broad non-monotonicity. For modulation rates of 480, 500, and 550 Hz thresholds obtained are within 5 dB of each other (-38.6, -34.4, and -36.9 respectively). It should be noted the ability to discriminate when the distance between the primaries is 800 Hz is uncommon. Thresholds between 650 and 800 Hz are relatively flat. Chance level performance was only found for the 400 Hz condition. Data for the 40 dB SPL condition are far more varied and the region of non-monotonicities for the 40 dB condition is somewhat hard to ascertain. There are three distinct non-monotonicities at 450, 550 and 600 Hz with thresholds as low -40 dB at 550 Hz. For intermediary frequency separations (500 and 560 Hz) thresholds are at chance levels.

According to the RQM, three minima between 400 and 650 Hz reflect a $3\pi$ (540 degrees) rotation of the CDT vectors, meaning that phase advances ~2.16 degrees/Hz. For the 40 dB condition, a fourth potential non-monotonicity is found for $f_m = 350$ Hz as there is a 10 dB decrease in thresholds between 280 and 350 Hz. Rate of phase change per degree is more difficult to estimate for the 80 dB condition as thresholds never return to chance levels above 400 Hz. We will estimate the start of the large non-monotonicity 400 to its minimum at 480 Hz as rotation of $\pi/2$ radians (90 degrees),
from which we estimate phase of the CDT advances 1.125 degrees/Hz. Unfortunately, due to limited resources and more data could not be collected.

Data for S2 (middle panel of Fig. 4.1) are more ambiguous. For an 80 dB SPL carrier, one sharp non-monotonicity was found for a frequency separation of 370 Hz and a smaller one at 450 Hz (thresholds being -22.4 and -9.1 dB, respectively) Although these non-monotonicities remain at 40 dB SPL carrier, a third non-monotonicity is present between 150 and 300 Hz with a minimum at 260 Hz. Thresholds are notably lower between 250 Hz and 300 Hz with a 40 dB SPL carrier than an 80 dB SPL carrier. A possible explanation reflects the phase of the CDT being in a detect state. Although $f_H/f_C$ is 1.1, less than the expected 1.22 for a maximal CDT, decreased thresholds may still represent a CDT affect (Hall, 1972).

Although S3’s carrier intensity was only reduced to 60 dB SPL, a clear intensity effect can be found. Rather than one minimum for modulation frequencies above 200 Hz, there are two. Again, and some thresholds are lower at lower intensities than at high intensities (such as at 400 Hz); however, this difference is small. Both S2 and S3 show a downshifting in the frequency separation that yields a minimum threshold when intensity is reduced. For S2 it is reduced from 500 to 425 Hz and from 370 Hz to 250 Hz for S2. While data has some points of ambiguity, predictions made from Hall’s experiment were supported. Overall, the number of non-monotonicities increases with decreasing intensity due to differences in phase. Phase differences from different intensities also effect the location of local minimum.
Chapter 5

Experiment 4: Testing the RQM and moving beyond AM and QFM

As seen in the previous chapters, most subjects’ threshold functions reach chance or near-chance performance as bandwidths increase before significant discrimination ability is regained. This occurs at larger bandwidths than would be plausible for a temporal modulation cue. If the tones are resolved independently, the mechanism behind discrimination at high modulation frequencies is not the output of a temporal filter, such as the filter imagined by Mathes and Miller (1947). If subjects are merely discerning between $\text{CDT}_{\text{am}}+f_L$ and $\text{CDT}_{\text{qfm}}+f_L$, a broader range of stimulus manipulations are open than could be considered previously as parameters of each component can be manipulated independently. As a starting point, we will consider one prediction of the RQM. The RQM predicts if the starting phase of the lower sideband were to be altered, thresholds would be affected. Assuming the presence of a perceptible CDT at the site of $f_L$, if the phase of $f_L$ were to be swept $2\pi$ (360 degrees), thresholds should reflect both non-detect and detect states twice.
5.1 **Subjects:**

Two subjects participated ages 20 and 29 with normal absolute pure tone thresholds for frequencies collected. Each received monetary compensation for their participation.

5.2 **Stimuli:**

Stimuli are similar to that in equation 1.1 except that $\theta_l$ is added to the low frequency component. $\theta$ of $f_c$ remains at 0 for the modified 'AM' stimuli and $\pi$ for modified 'QFM' stimuli.

$$y(t) = \sin(2\pi f_c t + \theta) + m/2 \sin(2\pi (f_c + f_m)t) + m/2 \sin(2\pi f_c - f_m)t + (\theta_l))].$$ (5.1)

5.3 **Design & procedure:**

Procedures were the same as described in the previous chapters. Subjects listened monaurally, with their right ears. Discrimination thresholds were measured with a 1-up 2-down adaptive procedure. Center frequency was again held constant at 2000 Hz and the amplitude of the center component was calibrated to 70 dB SPL. Thresholds were obtained from the standard AM/QFM discrimination task (as described in chapter 4) to determine the spacing between primaries that produces a minimum threshold at higher 'modulation frequencies' (450 Hz for S1 and 425 Hz for S2). Once an ideal frequency separation between components was established, the frequency of
components remained fixed, and the phase of the lower component, $\theta_l$, was introduced as an independent variable.

### 5.4 Results:

Data for both subjects are plotted below in Fig. 5.1 below.

![Figure 5.1](image)

Figure 5.1: Thresholds for S1 (shown in solid squares) and S2 (shown in open circles) are plotted as a function of phase of $f_L$. The dashed line at -2.65 indicates chance performance.

Data from both subjects support the hypothesis that low thresholds for high frequency separations are based on CDT interactions with $f_L$ and support the RQM. Although data for S1 (Fig. 5.1 solid squares) are limited in detail, thresholds clearly show the effect of the phase of $f_L$. A 10 dB improvement is found between $\theta_l = 0$ and $\theta_l = \pi$.
\[ \pi/6; \text{however between } \theta_l = \pi/6 \text{ and } \theta_l = \pi/4 \text{ thresholds rise by about 27 dB (from -35.9 to -9.2 dB), although performance always remains better than chance. It is unclear why thresholds jump at this point; however, as described previously, Hall (1972b), found some conditions with abrupt changes in phases for cancellation tones, where the phase of the cancellation tone (therefore, also the CDT) advanced like a step-function. A higher level of detail could not be gathered as resources were limited. Thresholds for } \theta_l = 0 \text{ and } \theta_l = \pi \text{ are nearly identical (-25.2 and -23.6 dB), the mirroring of thresholds when } \theta_l = 0 \text{ and } \theta_l = \pi, \text{ for the modified 'AM' and 'QFM' stimuli, lends strong support for the RQM.}

\]

Data for S2 (Fig. 5.1, open circles) tend to not have the dramatic rises and falls as shown for S1. However, it is also the case that S2 has higher overall thresholds for S2 (reaching a minimum of -16 dB (occurs at } \theta_l = 0 \text{ and } \pi ) \text{ than for S1. Data reflect a full rotation of } f_l \text{ relative to the CDTs, as predicted by the RQM, as S2 has two instances of being in a detect and non-detect state. Detect states occur for } -\pi/12 \text{ to } \pi/3; (\text{ -15 to 60 degrees}) \text{ and } \pi \text{ to } 7\pi/6; \text{ 180 to 210 degrees}); \text{ whereas, non-detect states occur between } \pi/2 \text{ and } 5\pi/6 \text{ (90 and 150 degrees) and between } 11\pi/6 \text{ to } 11\pi/6 \text{ (220 and 330 degrees).} \]
Chapter 6

Discussion

AM/QFM discrimination using tonal carriers offers a unique way to study temporal sensitivity as well as the perceptual effects of distortion products. Theorists have long hypothesized that measurements of temporal sensitivity with an AM/QFM discrimination task may be compromised by cues induced by distortion products. It was not predicted, however, that these effects would manifest as non-monotonic TMTFs, nor was it predicted that temporal and spectral cues would have different dominance regions with respect to modulation frequency. Tabuchi, Borucki, and Berg (2012) showed that thresholds at high modulation rates were degraded by randomizing the phase of the carrier, with little effect on thresholds at low modulation rates. It was also shown that thresholds at low modulation rates were degraded by decreasing the stimulus duration, with little effect on thresholds at high modulation rates. Perceptually, subjects reported listening to “pitch”, “heaviness”, or “loudness” differences at high modulation rates and “roughness” at low modulation rates.
The finding that temporal and spectral cues dominate at different regions of the TMTF increases the validity of estimates of temporal sensitivity. Although an overlapping region of saliency for both temporal and spectral cues might be substantial in some cases, as suggested by the data of S7 in Fig. 5.2, it can be inferred that the maximum bandwidth of phase sensitivity is in the range of 500 Hz to 600 Hz. A carrier frequency to bandwidth ratio of 3.3 is generally consistent with Tabuchi, Borucki and Berg (2012) for a 2 kHz carrier and with Bernstein and Oxenham (2006) for a 1 kHz carrier, but greater than ratios of 2.2 at 60 dB and 1.2 at 80 dB reported by Nelson (1994). The wider estimates of phase sensitivity reported by Nelson might be due to a difference in task, AM/QFM discrimination with 100% modulation depth, or to a difference in intensity, but it is notable that the psychometric functions for a number of Nelson’s listeners were non-monotonic, particularly at low intensities, suggesting the presence of a distortion product effect. Nelson concluded that distortion product induced differences in the envelopes of AM and QFM stimuli was the cause of the nonmonotonic data and claimed that the magnitude of CDTs could not provide enough energy to induce a level cue. If this were the case, the envelope cue should still be degraded with decreasing stimulus duration. If the range of non-monotonicities in these studies were to be misinterpreted and included as estimates of phase sensitivity, then the estimates would be comparable to those found by Nelson at higher stimulus intensities.
Although the archaic term ‘proximal’ stimulus is essentially synonymous with retinal image, the term is also appropriate for placing emphasis on the fact that displacements along the basilar membrane are a mixture of the exogenous or ‘distal’ stimulus and the endogenous effects such as distortion products, frequency-dependent phase delays, and the efferent system. Psychoacoustic models almost exclusively use the distal stimulus as input. In most instances, differences between the distal and proximal stimulus will likely be inconsequential, perhaps manifested as a source of internal noise in relation to optimal detection. In some instances, however, consideration of the proximal stimulus is critical. A case in point is that the explanation for the current findings depends on knowledge about the proximal stimulus, whereas no straightforward explanation follows from a consideration of the distal stimulus alone. The surprisingly large magnitudes of the distortion product effects reported here gives some concern that stimulus configurations from other experimental paradigms might be susceptible to ‘stimulus contamination’ when substantive differences exist between the proximal and distal stimulus, possibly leading to erroneous interpretations of data or imprecision of theoretical predictions.

Excepting the early psychophysical studies, most of the data about distortion products has been acquired through audiometric measurements, which also provide only inferential knowledge about the proximal stimulus. DPOAE measurements are subject to external factors such as stiffness of the ossicular chain or calibration and probe factors which can be difficult to normalize. Beyond external factors, measurements are subject to internal complications as well. DPOAE measurements are
typically thought of as the summed output from emissions generated from the
distortion source and those the wave scattering generated from preexisting mechanical
perturbations (Talmadge et al., 1999; Kalluri & Shera, 2001). To further complicate the
matter, measurements obtained are also believed to come from a variety of unknown
sources more broadly distributed on the basilar membrane that may be evident in
notches found in DP-grams (Martin et al., 2013). Taking observations at only at
discrimination ability at the site of the DP, no effort is required to unmix sources of
DPOAE output. Conceivably, a comprehensive understanding of the relationship
between DPOAE measurements and AM/QFM discrimination may yield an instructive
view of the proximal stimulus. What is currently lacking is a mapping of psychophysical
effects that is as complete as a DP-gram. Obtaining that level of detail with the probe-
tone, beat-cancellation technique would be painstakingly impossible, whereas the
current listening task would entail an effort no greater than that required for a standard
psychophysical discrimination experiment. Although the finding that two distinct
auditory processes underlie AM/QFM discrimination is important with respect to
psychophysical theory, the psychophysical technique per se will assuredly have greater
generality and value. Predictions for the RQM are generally supported by the data, and
in cases of disagreement, the model still provides a theoretical framework for
examining any discrepancies in greater detail.
References


