PRACTICAL APPLICATIONS OF DIFFERENCE TONES IN ELECTRONIC MUSIC COMPOSITION AND SYNTHESIS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MUSIC
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Alex Chechile
August 2020

This dissertation is online at: http://purl.stanford.edu/vr609df7082

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Chris Chafe, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jonathan Berger, Co-Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Paul DeMarinis

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Gerald Popelka

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

Difference tones are frequency components produced within the ear upon the physical and physiological interactions of spectral components in a given auditory input. Under certain conditions, these additional frequencies are audible and appear to be localized within the head. The phenomenon is the result of nonlinearities in the auditory system. The specific frequencies can be predicted by a classical power series expansion of a sum of sinusoids. With a two-tone signal, the quadratic and cubic terms of the polynomial yield two of the most audible difference tones.

Difference tones introduce creative opportunities when applied in musical contexts. However, the effect is highly dependent on parameters of the acoustic primary tones. The conditions for evoking difference tones with a two-tone stimulus are well understood, but a three-tone signal is more complex and not typically employed in music. For the effective implementation of three-tone difference tones in creative work, the parameters for reliable detection must be understood.

This dissertation consists of three interconnected branches of research. A psychoacoustic study on the detection of up to nine classes of two and three-tone difference tones reveals the most audible difference tones and the required stimulus conditions. Creation of an open-source toolbox of digital audio instruments for the synthesis of difference tones allows both creative and educational opportunities. With the results of the psychoacoustical study and the implementation of digital instruments, two series of compositions illustrate the physically of sound and spatial depth uniquely available with difference tones.

Acknowledgments

This dissertation is the product of an encouraged and supported curiosity instilled, inspired, and driven by mentors who I am fortunate to call friends and family. For always having time for thoughtful conversation, for invaluable support, and for his ability to make even the most mundane exciting and fun, I would like to thank my father, Richard Chechile. For teaching me the piano before my feet reached the pedals, attending a guitar convention that we *nearly* left empty-handed, and sharing enriching conversations about our most recent compositions, I would like to thank my mother, Jeannette Chechile. For the discussions about music, and for always rooting for me, I would like to thank my brother, Nicholas Chechile. I would also like to thank Audrey Hincman and Chris Hinson, Julia Hincman-Francavilla and Ben Francavilla, Zak Opaskar and Monica Guiterrez, and everyone in my extended family.

For the community he helped create, and for our friendship, I would like to thank John Chowning. Bookended by telematic collaborations in the mid-aughts and socially-distanced collaborations in the present year of 2020, I would like to express my sincere gratitude to my thesis advisor Chris Chafe for his guidance, support, inspiration, and friendship over the years. I would like to thank my other committee members Jonathan Berger, Paul DeMarinis, Brian Ferneyhough, and Gerald Popelka. Not only did my committee help shape this work, but I will be unpacking ideas we discussed for years to come. Additionally, I am grateful Terry Berlier was the chair of my oral defense committee.

For the collaborations we developed over the past seven years, I would like to thank Jonathan Abel, Dimitri Bouche, Jean Bresson, Constantin Basica, Elliot K. Canfield-Dafilou, David Fifield, Takako Fujioka, Mark Hertensteiner, Madeline Huberth, Loren

Mach, Romain Michon, Jérôme Nika, Spencer Salazar, Stephen Sansom, Steven Shepard, Suzanne Thorpe, Ge Wang, Kurt Werner, Bart Woodstrup, and Matt Wright. I would also like to thank faculty members in our department for the enriching conversations and support: Patricia Alessandrini, Mark Applebaum, Pierre Divenyi, Charlie Kronegold, François Rose, Malcolm Slaney, Julius Smith, and Erik Ulman. Jarek Kapuscinski deserves special recognition for his fearless guidance as a department chair, as well as for his ability to understand the core of my work with little explanation.

My experience at Stanford was greatly shaped by the individuals who have worked or are still working at CCRMA and in the music department, including Carlos A. Sanchez Garcia-Saavedra, Debbie Barney, Eoin Callery, Mario Champagne, Rob Hamilton, Christopher Jette, Dave Kerr, Sasha Leitman, Fernando Lopez-Lezcano, Velda Williams, and Nette Worthey. For making a profound first impression of the community at CCRMA, for his dedication to making internal and external musical connections, and for his warmth and humor, I honor the memory of Carr Wilkerson.

Between CCRMA, composition, and musicology programs, I would like to thank my fellow students for the community and friendship we share. This includes Yuval Adler, Myles Borins, Joe Cadagin, Michele Cheng, Luke Dahl, Orchi Das, Tysen Dauer, Hassan Estakhrian, Simona Fitcal, Nick Gang, Ethan Geller, Elena Georgieva, Francois Germain, Chet Gnegy, Emily Graber, Victoria Grace, John Granzow, Andrew Greenwald, Alex Hay, Holly Herndon, Julie Herndon, Ivan Naranjo, John Ivers, Megan Jurek, Nolan Lem, Shu Yu Lin, Jessie Marino, Sarah McCarthy, Doug McCausland, Tyler Mitchell, Prateek Murgai, Chryssie Nanou, Barbara Nerness, Cooper Newby, Mark Rau, Iran Roman, Nik Sawe, Charlie Sdraulig, Hana Shin, Laura Steenberge, Roshan Vid, Nick Virzi, Andrew Watts, and Julie Zhu. Additionally, I would like to thank the Stanford C2 program, and Steven Brill in particular.

For facilitating continued study, I am grateful for the opportunities provided by the Mellon Foundation and the Stanford School of H&S. For the experiences abroad, I would like to thank the Stanford Center at Peking University, IRCAM, and the Cité Internationale des Arts.

The work leading to this dissertation began far before my arrival at Stanford. I would like to thank my past mentors and friends including John McDonald, Don Berman, Paul Lehrman, Janet Schmalfeldt, Curtis Bahn, and Tomie Hahn. At Harvestworks, I would like to thank Hans Tammen and Carol Parkinson. I would also like to thank Éliane Radigue for her mentorship, enriching tea conversations, and friendship. Maryanne Amacher is largely responsible for setting me on the academic path of this dissertation topic. I am grateful beyond words for all of the late-night conversations we shared, and her guidance.

For teaching me that listening is boundless, I am forever indebted to Pauline Oliveros. I carry memories of her friendship, witty humor, visionary ideas, and iconoclastic thinking with me every day. Pauline's influence is deeply embedded in this dissertation.

I would also like to thank my friends, including Rebecca Autumn Sansom, Ethan Bach, Alon Ben-Shoshan, Olga Berar, Lauren Blasingame-Pack Ben-Shoshan, Mike Bullock, Maria Chavez, Meredith Corey, Kate Daly, Natalie Deam, Brandon Farris, the Gerrein Family, Adam Ghory, Gary Graham, Tim Grimes, Ione, Lee Todd Lacks, Wil Lindsay, Heidi Manschreck, Noah Marchal, Aaron Martin, Brian Moran, Will Murray, Hugo Paris, Carl Robinson, Sam Rosen, Amanda Schupak, Micah Silver, Chris Skinner, Kate Smolik, Sergei Tcherepnin, Stefan Tcherepnin, Doug Van Nort, and Charles Veasey.

Finally, I would like to express my deep gratitude to Cathleen Grado for always listening, advising, and being a balancing force throughout this adventure.

Preface

My first experience with difference tones dates to a 2003 rehearsal of a piece I wrote while at Tufts University. As I was walking around an open grand piano listening to John McDonald perform *Prelude for Music for Eleven Stereos*, I noticed ghostly tones that were not emerging from the piano, but from somewhere else entirely. Shortly after, I came upon the work of Maryanne Amacher, and her deliberate use of sound to elicit physical and neural responses resonated deeply. While attending Rensselaer Polytechnic Institute a few years later, my friend and colleague Micah Silver introduced Maryanne and me, and thereafter we developed a close friendship. We would spend hours discussing ideas that teetered between science fiction and possibilities for future works. Usually these conversations stemmed from articles in Scientific American, Nature, or other online publications on science and technology. Around this time, I would experiment with difference tones as an intentional layer of sound using my modular synthesizer. After Maryanne passed, I began composing works with the phenomenon in a way of memorializing her, expanding ideas from our conversations, and connecting threads back to my work at Tufts. My first public performance using difference tones was in 2010 while opening for the band Primus in an at-capacity venue of 2,700 people. The speaker system was enormous, and I hope the performance unexpectedly opened some ears like Maryanne opened mine.

The content of this dissertation is presented in a linear narrative (background, model, psychoacoustic study, software, and music), but the work was conducted in an order entirely of its own. Composing with difference tones was a constant throughout the process, as well as designing musical instruments to facilitate the works. Literature reviews continuously informed my practice and shaped the pieces and instruments as they developed. The psychoacoustic study occurred towards the end of the process, and a resulting series of etudes emerged from exploring the data in music. In other words, my understanding of the effective use of difference tones developed through writing a series of pieces, and questions that arose from the compositional process drove the development of further research. The individual elements are detailed in respective chapters, but like the phenomenon itself, the development of the dissertation was also a nonlinear process.

The first chapter addresses questions pertaining to the benefits of difference tones in creative practices and outlines the dissertation's main contributions. The chapter provides the background history of the phenomenon, the biomechanics involved, and covers related subjects such as linearity, nonlinearity, and other perceptual effects. Concluding the chapter is an overview of artists and past works that use difference tones in both music and sound installation.

The second chapter provides a model for calculating distortion product frequencies from multi-component stimulus. Examples for two and three-tone primaries show how the method yields the general frequency relationships of the quadratic and cubic distortion products. A software implementation calculates the results for any number of frequency components, provides the frequency relationships, gives plots for visualization, and supplies other helpful information for perceptual analysis.

The third chapter describes a psychoacoustic study investigating the perception of up to nine classes of two and three-tone difference tones across a wide frequency range. The results form a database showing the required stimulus parameters for reliably evoking various classes of quadratic and cubic difference tones. The chapter also provides an overview of methods for measuring difference tones, results from past perceptual studies using cancellation techniques, and related topics.

The fourth chapter outlines the Ear Tone Toolbox, a collection of open-source audio instruments that evoke difference tones with precision. Details on each of the five main unit generators are provided while simultaneously discussing the various software and hardware formats available. The underlying algebraic relationships among acoustic primary tones and the difference tones are useful for learning the fundamentals behind the phenomenon.

The fifth chapter discusses the unique characteristics and benefits of difference tones in music, as well as best practices. A series of works entitled *On the Sensations of Tone* functions as case studies demonstrating the effective use of difference tones in larger compositional contexts. A second series called *The Ear Tone Etudes* illustrates particular findings from the psychoacoustic study of Chapter 3. The chapter ends with a discussion on peripheral works also involving difference tones.

Summarizing the topics covered in the dissertation, the sixth chapter is a postscript that reflects on the interdisciplinary nature of the work as a whole and suggests future directions. It is followed by appendices containing code for a mathematical model, and musical scores.

Table of Contents

List of Tables

List of Figures

CHAPTER 1
Historical Background, Questions, and Groundwork

1.1 INTRODUCTION, DEFINITIONS, AND TERMINOLOGY

The auditory system is responsible for the sensation and perception of sound, but what is possible if the system itself is directly manipulated to become an active component in the production of sound? How can this phenomenon change our relationship to sound, and how we experience music, installation, and sound art? With our current understanding of auditory perception, and through the use of current technology, it is possible to answer these questions and push the boundaries of immersion in the sound-based arts through the development of new instruments, synthesis techniques, and theoretical applications.

Auditory distortion products (DPs) are tones that are generated in the inner ear in response to acoustically presented primary tones. The frequencies evoked in the ear, which are also known as intermodulation components, are the result of nonlinearities of the auditory system, and are primarily generated by the outer hair cells in the cochlea. The DPs are sensed directly in the inner ear, but the frequencies also propagate backwards from the inner ear, through the middle ear, and out to the ear canal where they can be recorded by an in-ear microphone. These backwards propagating sound waves are called distortion product otoacoustic emissions (DPOAEs), and are used clinically to test aspects of cochlear functioning.

Auditory distortion products are created along the basilar membrane in the cochlea in response to the simultaneous presentation of at least two frequency components f_1 and f_2 ($f_2 > f_1$). The combination of these frequencies produces multiple distortion products, but the two most audible are at the frequency f_2-f_1, which is called the quadratic distortion product (QDP) or the quadratic difference tone (QDT), and at the frequency $2f_1$-f_2, which is called the cubic distortion product (CDP) or the cubic difference tone (CDT). Other distortion products are also produced, but not all are audible. As discussed in Chapter 3, the specific conditions for audibility of distortion products often depends on a number of interconnected criteria, such as the overall amplitude, frequency range, and frequency separation between the external primary tones. Additionally, the various difference tone classes usually exhibit their own relationships among these parameters. For example, the CDT is more sensitive to the frequency ratio between the primary tones than the QDT.

The discovery of additional tones not present in the external acoustic tones is attributed to the composer and violinist Giuseppe Tartini who noted the effect when playing double-stops on his violin (Tartini, 1754). In tribute, DPs are sometimes called Tartini tones in musical contexts (typically when evoked by non-sinusoidal stimulus frequencies, such as a violin or other acoustic instruments). Because the effect is produced by the simultaneous sounding of two tones, the term combination tones was coined by Vieth (1805), and later popularized by Hermann von Helmholtz, who preferred the term since it includes both the aforementioned difference tones as well as summation tones f_1+f_2 (Helmholtz, 1863). However, summation tones generated by auditory system nonlinearities are rarely audible. I will use the more general terminology auditory distortion products and combination tones interchangeably, and use summation tones or difference tones when referring to either specific class.

1.2 MAIN QUESTIONS

While the phenomenon of auditory distortion products has been studied since Tartini, our understanding of the biomechanics involved is still not fully established, and new research continues to offer increased insight. With recent advancements come opportunities to re-evaluate how DPs function in music, to develop new synthesis techniques, and to advance methods and techniques in presentation. The precision of digital signal processing, increased linearity and fidelity of speaker systems, and evolution of technology brings opportunity to reassess how combination tones function in contemporary music and the arts. Additionally, digital synthesis techniques can incorporate perceptual research to evoke a more concentrated effect. By advancing our perceptual understanding of auditory distortion products, we move towards creating new experiences in music and other sound-based works.

Practical theories on the benefits of auditory distortion products in music require an update. After conducting investigations on distortion product detectability thresholds, Plomp (1966) wrote, "All mean detectability thresholds investigated in our experiments exceed 40 dB, corresponding with a nonlinear distortion of below 1%. As Janovsky (1929) and von Braunmühl and Weber (1937) have demonstrated, 3 to 5% distortion of speech and music can be introduced without being noticed, so we may conclude that, for usual listening levels, the ear's distortion is sufficiently low to avoid audible combination tones. This fact makes it rather improbable that combination tones represent a constitutive basis for musical consonance, as was stated by Krueger (1906-1910) and Hindemith (1940)." Since the time of Plomp's investigations, the fields of electronic music, sound art, and installation renders his conclusion outdated. The theories of Krueger and Hindemith (which will be addressed later in this chapter) can benefit from an update, particularly in regard to non-tonal music.

Most compositional and perceptual work involving auditory distortion products focus on the $2f_1$-f_2 CDT and/or the f_2-f_1 QDT with a two-tone signal. Several more difference tones are available with the addition of a third frequency component in the external signal, yet relatively few psychoacoustical studies investigate the perception of three-tone distortion products, and generally they are not used in music. To successfully employ three-tone difference tones in creative work (and eventually >3 tones), the parameters for reliable detection are needed. Specifically, what are the required ratio boundaries between acoustic primary tones, how does their behavior change across the frequency spectrum, and which classes of difference tones are the most audible? A psychoacoustic study investigating the detection of difference tones resulting from three-tone stimulus will answer these questions.

The impetus for developing new psychoacoustical work on the perception of three-component difference tones stems from composing with the phenomenon. While it is possible to evoke difference tones with acoustic instruments, the effect is more concentrated and precise when restricting destructive interference in both the acoustic space as well as within the ear. Due to the absence of harmonics, sinusoidal signals are ideal for evoking difference tones, however the resulting acoustic timbre is significantly simplified (by design). By adding additional sinusoids to evoke an even more complex set of difference tones, spectral complexity increases in both the acoustic space and the inner perceptual space. How can we effectively create digital instruments for auditory distortion product synthesis using an increased spectrum of primary tones, thus generating a complex distortion product spectrum? Of particular interest are parameters and techniques for producing music with non-tonal and frequency-independent complexes of sinusoids.

The first step in approaching these research questions involves a simple mathematical model that predicts the quadratic and cubic difference tone with any number of primary tone components (≥ 2). The model produces a large number of potential DPs, however, by taking past research into account, it is possible to restrict the number of components to those most-likely to be detectable. For example, summation tones have shown to be difficult to perceive, hence they can be disregarded. Using the model in parallel with software instruments allows for experiments in composition and perception.

The model, software, and findings from the psychoacoustic study are put to use in a series of works that exhibit the unique properties of distortion product synthesis. The techniques employed in each of the compositions forms a practical approach for evoking distortion products in contemporary music. Topics covered include expanded spatial dimension, increased harmonic content, modes of listening, movement and reactivity, among other subjects.

1.3 CONTRIBUTIONS

The scope of this dissertation encompasses three main components: a psychoacoustical study investigating the perception of difference tones with two and three-tone stimuli, digital audio instruments that evoke specific difference tones with precision and clarity, and a series of compositions that create a complex spectrum of difference tones. Before addressing these branches of interconnected research, a history of relevant topics centering around distortion products provides context, and a mathematical approximation of distortion product frequency components serves as a simple model. The results of the psychoacoustic investigation inform the development of the software, which in turn is employed in the compositions. As the research and instruments developed from questions that arose during composition, the three areas of research are intimately connected. Using the compositions as case studies, a discussion on the integration of difference tones in electronic and electroacoustic music provides novel methods for the effective use of difference tones in contemporary music.

1.4 BACKGROUND

The remainder of this chapter provides an overview of the principles behind the generation of auditory distortion products, a history of early experiments investigating the phenomenon, and notable examples of difference tones in music and sound art. Before discussing the biomechanics of the auditory system, we will first begin by discussing the fundamentals of mechanical and electrical linear and nonlinear systems, and intermodulation distortion in general.

A system is said to be linear if the output change is proportional to the input change, without any other distortion. For example, if we input a signal into a high-fidelity sound system and changed the amplitude (multiplication by a constant), the resulting output amplitude would change identically. This property is called homogeneity and must be present for a system to be linear. In mathematical terms, if we have a system with an input signal of $x(t)$ that produces an output of $y(t)$ and we scale it by a constant $ax(t)$, the output would be $ay(t)$. The second property of a linear system is additivity, meaning if two or more driving forces are simultaneously applied to a linear system, the output response would be the sum of the outputs if the inputs were applied individually. Referring again to our sound system example, additivity means that one component frequency would not perturb the other, as they coexist peacefully without any distortion in the system. Mathematically, additivity with two inputs $x_1(t)+x_2(t)$ would result in an output of $y_1(t)+y_2(t)$.

In contrast, the output of a nonlinear system will produce frequencies that are not present in the input, making the output non-proportional to the input. If a single sinusoid is applied to a nonlinear system, the resulting output signal will contain harmonic distortion defined as integer multiples of the single input frequency. Two or more different sinusoids simultaneously applied will produce intermodulation distortion between the input frequencies. The components introduced by the intermodulation distortion are at the sum (f_1+f_2) and difference (f_2-f_1) frequencies of the input signals f_1 and f_2, as well as the harmonics mentioned previously. There will also be intermodulation products between the interacting harmonics and fundamentals, producing additional frequencies such as $2f_1$-f_2, $2f_2$-f_1, $2f_1$+f_2, $2f_2$+f_1, $2f_2$+$2f_1$, $2f_2$-$2f_1$, etc. The amplitudes of the distortion products will vary, but they will be less than the amplitudes of the f_1 and f_2 input frequencies.

The nonlinearity is due to the multiplication of the two signals, hence the use of the word *product* in the term distortion product. In general, the multiplication between two signals is an example of nonlinear systems, and can include squaring a signal, multiplying two different signals, etc. For example, with the simplified equation:

$$\mathrm{R}(t) = \mathrm{Sin}(\omega_1 t) + \mathrm{Sin}(\omega_2 t) + \mathrm{Sin}(\omega_1 t)\mathrm{Sin}(\omega_2 t)$$

the first two terms are the linear components of the two frequency inputs, and the third term, the multiplication between the two, is the nonlinear term. After employing a few trigonometric identities (to be addressed in Chapter 2), and substituting ω with f, we are left with the frequencies f_1, f_2, f_2-f_1 and f_2+f_1, otherwise known as the quadratic distortion products. Similarly, the following simplified equation:

$$\mathrm{R}(t) = \mathrm{Sin}(\omega_1 t) + \mathrm{Sin}(\omega_2 t) + \mathrm{Sin}^2(\omega_1 t)\mathrm{Sin}(\omega_2 t)$$

would yield the frequencies f_1, f_2, $2f_1$-f_2 and $2f_1$+f_2, the cubic distortion products. At this point, we can see that the intermodulation products are in fact the same as the distortion products produced in the ear. The reason behind the nonlinearities generated in the ear will be discussed shortly.

In electronics, and electronic music, the multiplication between two signals is the basis for certain effects and synthesis techniques such as amplitude modulation and ring modulation, each of which produces sideband frequency components. Amplitude modulation is simply:

$$A(t) = (M(t) + 1.0) \times C(t)$$

where M is a modulator signal and C is the carrier signal. In other words, amplitude modulation involves the multiplication between a unipolar modulator and a bipolar carrier. It produces sidebands at the sum and difference frequencies (at half the amplitude of the carrier signal) for every sinusoidal component in the carrier and modulator frequency when the modulator is above 20 Hz. For example, with a 1000 Hz carrier and a 300 Hz modulator, the resulting signal will contain three components: 1000 Hz, and sidebands at 700 Hz and 1300 Hz (at half the amplitude). Ring modulation is similar, but both multiplied signals are bipolar and the carrier signal is not present in the output. It is shown by:

$$R(t) = M(t) \times C(t)$$

where M is the modulator signal and C is the carrier signal. When the modulator is above 20 Hz, the resulting signal changes timbre because the modulator generates a pair of sidebands for every sinusoidal component in the carrier, and the carrier frequency is not preserved. For example, with a single carrier at 1000 Hz and a single modulator at 300 Hz, the result of the ring modulation would yield only two sidebands at the frequencies of the sum tone and difference tone of the carrier and modulator signals: 700 Hz and 1300 Hz. The intermodulation products generated by these two synthesis techniques is related to the phenomenon of auditory distortion products.

In the above description of multiplication of two signals, we saw the simple difference tone f_2-f_1, and the $2f_1$-f_2 CDT arise from two sinusoids interacting in a nonlinear system. Because the auditory system contains several nonlinearities, it should not be surprising that these frequency components can also be generated in our ears. So what and where are the specific nonlinearities that generate auditory distortion products?

1.5 BIOMECHANICS

When sound enters the cochlea, the basilar membrane acts as a transducer conveying the sound vibrations in the cochlear fluids to the sensory inner hair cells (IHCs), which produce electrical signals that are relayed to the auditory brainstem through the auditory nerve (afferent innervation). At the same time, the outer hair cells (OHCs) receive electrical signals from the brainstem and mechanically vibrate at the frequencies of the sound (efferent innervation) (Brownell, 1990). This electromotility mechanically increases stimulus-specific vibrations on the basilar membrane, resulting in an increase of hearing sensitivity and frequency selectivity when transmitted to the IHCs (Gold, 1948; Davis, 1983). However, the OHC movement does not occur exclusively at the stimulus frequencies, but is somewhat irregular, making its frequency response nonlinear, extending to an audible range (Brownell, 1990). This nonlinear active process increases basilar membrane movement, while an excess of the generated energy causes additional vibrations that travel back from the basilar membrane to the middle ear and the ear canal to create OAEs (Kemp, 1978, 2003). While OAEs can be recorded directly in the ear canal with a specially designed earpiece, auditory distortion products are specifically the intermodulation components in the inner ear, and under certain conditions they can be perceived by the subject.

More specifically, the primary source of nonlinearity is associated with the outer hair cell mechanoelectrical transduction (MET), which is the process of transduction between the mechanical vibrations to electrochemical signals. When the OHC is stimulated by mechanical vibrations along the basilar membrane and thus the organ of Corti, a selection of individual bundles of stereocilia on the hair cell body are deflected. The stereocilia bundles on a cell are connected by tip-links, which are coupled to ion channels. When mechanically opened through stereocilia deflection, the channels let positive ions such as potassium and calcium enter the cell, which triggers the release of neurotransmitters in the cell resulting in an electrical nerve signal (Heller, 2013). The ion channels open when the tip-links are pulled, and close when they are pushed, but this process is nonlinear as the links can buckle, resulting in a disproportionate electrical potential output to the input sound pressure (Ashmore, 2000).

1.6 EARLY HISTORICAL REVIEW

Our current understanding of auditory distortion products has been continuously developing since the first discoveries of the phenomenon by Italian violinist and composer Tartini, French savant Romieu, and German organist Sorge in the eighteenth century. It was Sorge (1744) who first wrote about the phenomenon, followed by Romieu (1751) and Tartini (1754), but it is accepted that Tartini was the first to discover the phenomenon for he dated his discovery to 1714, and he taught the technique at his school for violin in Padua for 26 years (Boring, 1942). Since Sorge did not claim to discover the phenomenon in his writing, it is likely that he was reporting on Tartini's discoveries from interactions with students who studied with him (Boring, 1942). It is interesting to note that in these three early documents, the frequencies reported by Tartini, Romieu and Sorge do not completely match. The following early studies on the perception of difference tones were collected in a seminal review by Plomp (1965, 1966), and are summarized in this section.

In 1827, the French baron Blein conducted experiments using a violin to find the QDT and/or CDT for each interval within an octave, with the results plotted in Figure 1.1 (Weber, 1829). Hällström, also using a violin, proposed that additional tones consisting of interactions among the difference tones and the primary tones, as well as among the difference tones themselves, may give rise to additional combination tones, although he was able to only perceive one or two of the proposed four in Table 1.1 (Hällström, 1832). It is interesting to note how Hällström viewed the generation of the second, third, and fourth combination tones. The second combination tone, which we know as the CDT, relied on the interaction between the f_1 primary tone and his first combination tone,

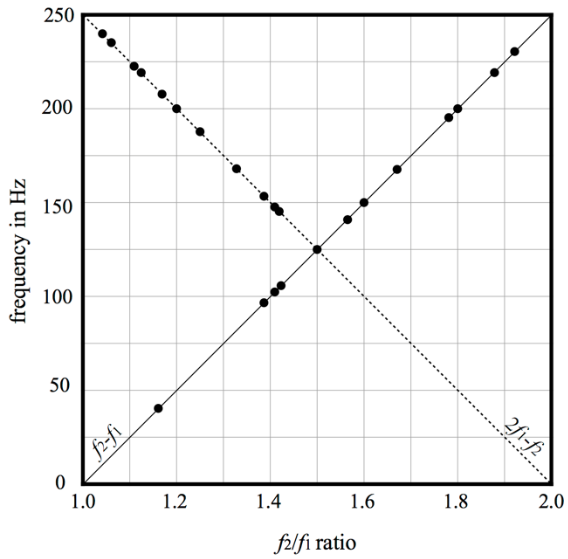


Figure 1.1. Replotted results from Blein's experiments showing detected QDT (solid line) and CDT (dashed line) with a f1 of 256 Hz, as reported by Weber (1829) and originally plotted Plomp (1965). The solid and dashed lines indicate the trajectory of possible tones per DT class, and the dots specify perceived distortion products.

CT #	First Component	Second Component	Combination Tone
I	f_2	f_1	$CT_1 = f_2 - f_1$
II	f_1	$f_2 - f_1$	$CT_2 = 2f_1 - f_2$
III	f_2	$2f_1 - f_2$	$CT_3 = 2(f_2 - f_1)$
IV	$2f_1 - f_2$	$f_2 - f_1$	$CT_4 = 3f_1 - 2f_2$

Table 1.1 Hällström's proposed tones, with converted notation of f_1 and f_2 ($f_2 > f_1$).

which we know as the QDT. The third combination tone was also due to the interaction between one of the primary tones, this time f_2 and his second combination tone, which again we know as the CDT. His fourth combination tone was the result of two other combination tones alone, the difference between the CDT and the QDT, giving rise to a combination tone at $3f_1$-$2f_2$. Figure 1.2 replots the results of Hällström's experiments, showing a larger number of combination tone classes. These additional combination tones were later criticized by Ohm (1839), who claimed they were all first order combination tones that were the result of interactions between the harmonics of the primary tones.

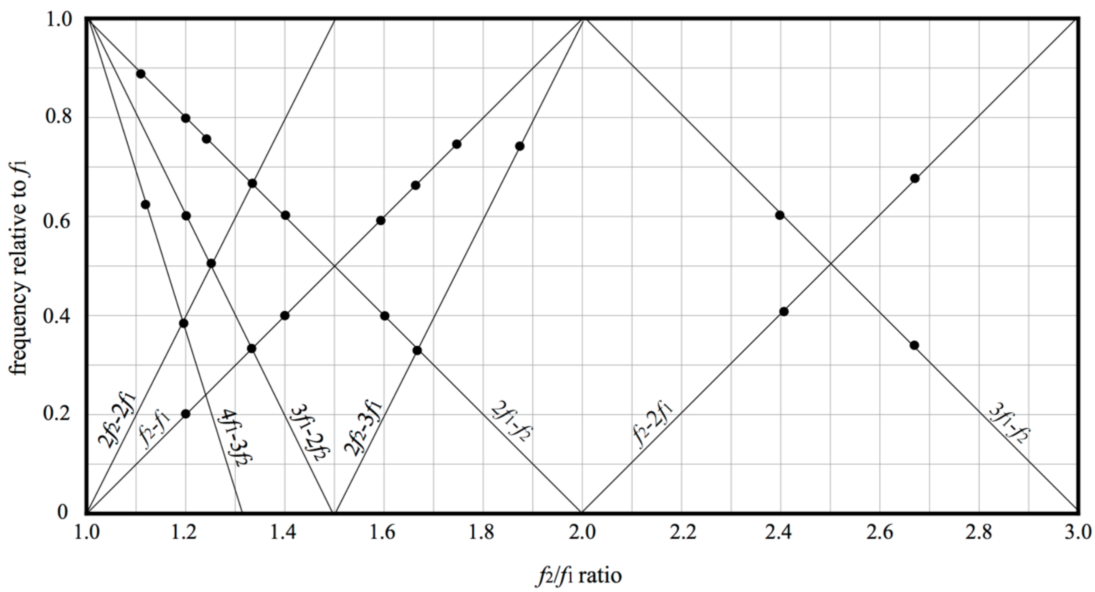


Figure 1.2. Results from Hällström (1832) displaying an increased number of combination tones with f_1 between 512 Hz and 1440 Hz. The replotted data was originally visualized by Plomp (1965).

Helmholtz investigated combination tones further and contributed the first model of the nonlinearity (Helmholtz, 1856, 1863). Using pure tones, he detected the f_2-f_1 QDT, and with non-sinusoidal tones he detected the $2f_1$-f_2 CDT. Helmholtz confirmed the tones that Hällström described, and reported on the number of tones heard per interval: zero for the octave, one for the fifth, two for the fourth, three for the major third and major sixth, four for the minor third, and six for the minor sixth. Helmholtz was also the first to identify that the perceived tones were due to nonlinear distortion of a mechanism inside the ear, much like mechanical nonlinearities.

König (1876) investigated the primary tone ratio using tuning forks and resonator boxes. He found both the QDT and CDT were detectable when the ratio between f_2/f_1 was less than 2.0 (an octave). For intervals forming a ratio between 2.0 and 3.0, König reported the f_2-$2f_1$ and $3f_1$-f_2 DTs. For intervals with a ratio between 3.0 and 4.0, he reported tones at the frequencies of f_2-$3f_1$ and $4f_1$-f_2.

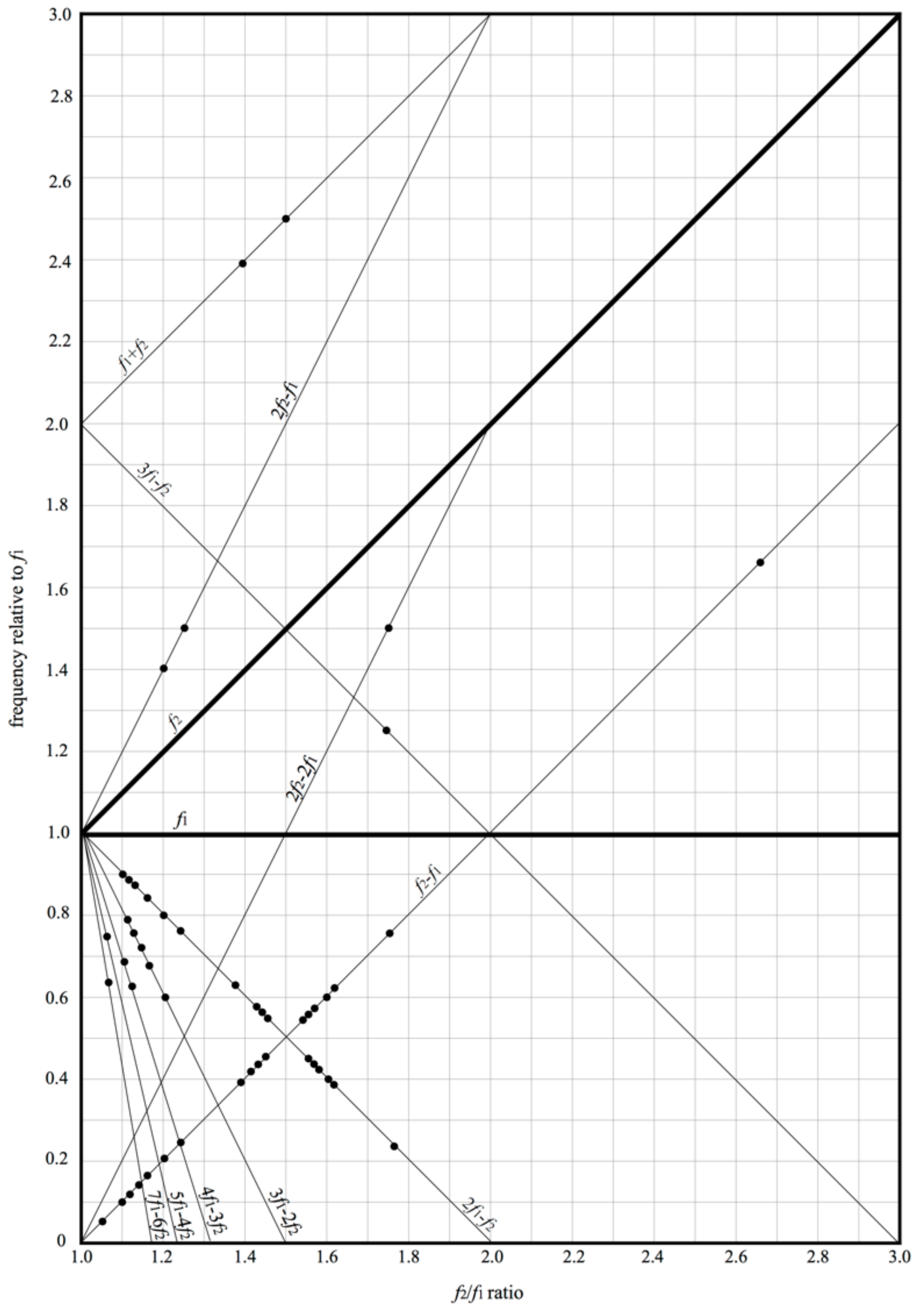


Figure 1.3. Results from Meyer (1896) showing combination tones resulting from a f_2/f_1 ratio of up to 3.0. Replotted from a figure in Plomp (1965).

Also using tuning forks and resonator boxes, Meyer (1896) found that the CTs between the f_1 and f_2 primary tones were difficult to perceive, and the CDT was louder than the QDT. Figure 1.3 reports Meyer's findings with a range of f_1 frequencies between 400 Hz and 1600 Hz. The figure displays the f_1 and f_2 primary tones as bold horizontal and 45 degree lines, with the detected combination tones shown as points along the lines at other angles.

As recounted by Plomp (1965), Krueger conducted thorough experiments that resulted in a theory of musical consonance based on combination tones (Krueger 1900-1910). Key points in his theory were 1) larger intervals yielded fainter combination tones, 2) for f_2/f_1 less than 2.0, the QDT and CDT were louder than his "higher order" combination tones, 3) the combination tones between the primary tones were more difficult to perceive, 4) the intervals that shared equivalent CTs produced the loudest CTs, and 5) the summation tone was not always perceptible. Krueger's experiments were replicated by Stumpf (1910), who rejected the theory as he was only able to recreate the QDT and CDT over a wide range, therefore excluding Krueger's "higher order" combination tones.

Using sine tone signal generators, Plomp (1965) conducted a series of experiments investigating detectability thresholds for CTs after noting that past work reported different CTs beyond the QDT and CDT. Plomp reasoned the previous incongruities were due to the inability to control stimulus tone amplitude. The results showed that higher stimulus decibel levels led to the detection of more CTs, and the QDT, CDT, and $3f_1$-$2f_2$ CTs were the only tones detected by all subjects. Other tones were found, but varied based on the individual. Also, Plomp found the ratio between primary tones that yielded the most CTs was up to 1.5, or a musical fifth.

An era involving more advanced methods of psychoacoustic difference tone measurement developed around this period. Up to this point, the studies described in this section used a method where the subject (usually the investigator) matched the pitch of the difference tone with a variety of sound generating instruments. A golden age of distortion product measurement began with Zwicker (1955) and Goldstein (1967), and the era continued primarily until the next paradigm shift with the discovery of DPOAEs in the late 1970s. The era of psychoacoustic measuring of DTs is described in Chapter 3, and the next section traces the history of DPOAEs.

1.7 THE PATH TO DPOAE

The following three paragraphs were adapted from a conference paper presented at the International Conference on Auditory Display (ICAD) (Chechile, 2015).

Contrary to Békésy's widely accepted theory that the cochlea acted passively, British auditory scientist Thomas Gold believed an active electro-mechanical process was needed to counteract the damping that occurs through the viscosity of the liquid in the inner ear (Gold, 1948). His ideas were not embraced at the time, however in the mid-1970s the British physicist David T. Kemp also thought that a mechanically active process was taking place in the ear. By placing a microphone in the ear canal, Kemp was able to capture sound produced in the external ear canal, making the discovery of what we now call spontaneous and distortion product otoacoustic emissions (Kemp, 1978, 2003).

Physiologist Hallowell Davis (1983) described a "cochlear amplifier" involving an active process with the outer hair cells. First observed by biophysicist William E. Brownell, the OHCs on the basilar membrane of the cochlea are capable of electromotility (as described earlier in this chapter), which is the most significant mechanism for the generation of otoacoustic emissions (Brownell, 1990). Hall (2000) concisely states, "stimulus-induced macromechanical vibrations of basilar membrane and organ of Corti produce micromechanical vibrations of hair cells. These can in turn exert a 'feedback' influence on the macromechanics." The frequency response of the electromotility is nonlinear and extends to an audible range (Brownell, 1990).

Kemp (2008) explains that otoacoustic emissions are the result of a leakage of energy from the cochlear amplifier. He states, "in reality, as outer hair cell gain is increased, there comes a point where any small irregularities in outer hair cell arrangement activity become magnified and significant stimulus frequency energy travels backward, to cause OAEs" (Kemp, 2003). Otoacoustic emissions testing is now widely used as an objective physiologic measure of cochlear functioning, and a particularly effective method for testing newborn hearing health (Baldwin *et al.*, 2010; Dhar & Hall, 2012).

The distortion product primarily investigated using DPOAEs is the CDT $2f_1$-f_2, and the parameters for stimulus ratio, amplitude, and frequency range have been studied by a number of researchers (Harris *et al.*, 1989; Gaskill & Brown, 1990; Hauser & Probst, 1991; Popelka *et al.*, 1993; Rasmussen *et al.*, 1993; Nielsen *et al.*, 1993; Kemp 1998). Typically, a f_2/f_1 ratio that yields maximum DPOAE levels is around 1.22, with primary levels either varied (L_1>L_2) or equal (L_1=L_2), and the response varies depending on the range of stimulus amplitude levels. As will be discussed in Chapter 3, distortion product levels measured by DPOAE and psychoacoustic methods do not typically agree.

1.8 PERIPHERAL PHENOMENON

The missing fundamental effect is a phenomenon where the fundamental frequency of a complex tone can be perceived despite its spectral absence (Seebeck, 1841; Schouten, 1938, 1940; Licklider, 1951; de Boer, 1956). For example, if we have a tone consisting of three sinusoids at 600 Hz, 700 Hz, and 800 Hz, the missing fundamental would be the phantom pitch that could be perceived at 100 Hz, a tone that is not physically present. Also known as the residue pitch, virtual pitch, or periodicity pitch, the perception of the fundamental pitch relates to neural processing of repetition patterns of higher integer multiples. Although there are exceptions (e.g., an existence region), the pitch of the missing fundamental generally can be estimated by finding the greatest common divisor of the component frequencies.

Early theories incorrectly attributed the missing fundamental to difference tones. Considering our example of a signal with 600 Hz, 700 Hz, and 800 Hz tones, the 100 Hz missing fundamental is the same frequency as the QDT of any two consecutive components. However, the two phenomena were proven to be separate through cancellation and pitch-shifting experiments (Schouten 1938, 1940). Additionally, auditory distortion products require two or more frequency components in a single cochlea, but the missing fundamental can be observed with one tone presented to one cochlea and the other tone presented to the opposite cochlea (Houtsma & Goldstein, 1972).

Even though the two phenomena have been shown to be separate, distortion products can contribute to the perception of the missing fundamental. Pressnitzer & Patterson (2001) conducted a series of experiments using harmonic primary tones consisting of sinusoids spaced by 100 Hz to investigate a "distortion spectrum" (DS) of QDTs. With two subjects, they found the first four components of the DS were above the hearing threshold, the DS could be evoked by primary tones at low to moderate loudness levels, and the 100 Hz CT at f_0 (the frequency of both the missing fundamental and the QDT) had an amplitude that increased when the number of primary tones in the harmonic tone complex was increased. As the stimulus in both the Pressnitzer & Patterson and the missing fundamental studies are harmonic signals without the lowest components, this study suggests it is possible for a harmonic signal to produce strong distortion products that could contribute to the perception of the missing fundamental. We will revisit this study in subsequent chapters.

1.9 DIFFERENCE TONES IN MUSIC

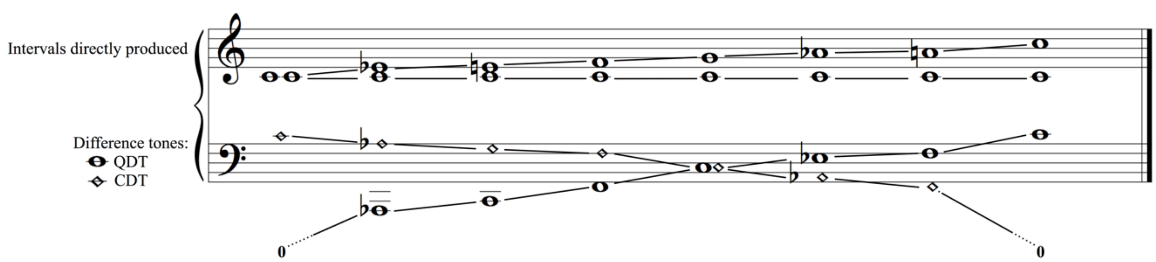


Figure 1.4. Hindemith's (1945) notation of QDT and CDT from intervals.

Approaching the mid-twentieth century, Hindemith (1945) dedicated a significant portion of text in the first volume of *Craft of Musical Composition* to his theory of musical consonance based on combination tones, which he describes as "important for the subconscious ear." He recognized the advantages that mid-twentieth century technology offered to the study of CTs when compared to the previous work conducted with acoustic instruments. His analysis encompassed the CTs produced between note intervals, and the relationship between CTs and the harmonic series. In his language, he described the QDT as "first order" and the CDT as "second order" combination tones. His rationale behind the generation of the "second order" (and beyond) was similar to Hällström, and provided the example deriving the CDT from a minor dyad: the 320 Hz E4 and 384 Hz G4 produced a 64 Hz C2, which was his "first order," also known as the QDT. The E4 and the C2 CT interacted to form a 256 Hz "second order" CDT. He also reported the CT between G4 and C2 produced a note at one of the primary tones, the E4, so it was not considered beyond intensifying the primary tone. This example showed his musical way of thinking about the phenomenon otherwise understood through the mathematics of a nonlinear system. Figure 1.4 shows Hindemith's notation of intervals and the resulting quadratic and cubic difference tones.

Hindemith discusses CTs as representing a "clouding or a burdening of the interval," referring to the strengthening and weakening effects on the consonance of each interval. Specifically, the strengthening and weakening was based around a theory involving the relationship between the stronger QDT and weaker CDT and the position of the "root" note of the dyad. In regard to intervals excluding the unison, octave, and the perfect fifth, Hindemith states, "all other intervals carry a double burden of varying weight. The clouding of the interval is not so strong that any effort need be made to suppress the combination tones. Yet of course they must not be so strong as to overshadow the directly produced interval." As discussed by Plomp (1966), the typical level of unamplified music may be insufficient for drawing these conclusions, but the possibility DTs have a subtle effect is valid (hence Hindemith's "subconscious ear" comment). Difference tones can be evoked with increased intensity using amplified electronic instruments. Just as Hindemith acknowledged the benefit of electronic instruments in the study of combination tones, the same instruments are used in music, which brings weight to his theories of consonance.

1.10 EXAMPLES IN COMPOSITION AND INSTALLATION

While musicians have been aware of the phenomenon since Tartini's teachings, specific cases of CTs in composition prior to the 20^{th} century are not discussed as frequently. Campbell and Greated (1987) described an example of perceived difference tones in the finale of Sibelius' *Symphony No. 1* in E minor, Opus 39. In a brief passage, a pair of flutes enter unmasked by other instruments and evoke QDT and CDT by playing in thirds at a *mezzo-forte* dynamic. This case was detected while one of the authors was participating in a rehearsal of the composition, but it may be difficult to perceive from a recording.

Deliberate examples emerge in the work of Maryanne Amacher, whose use of electronics to evoke the phenomenon extended the field to our contemporary practice. The ear tones in Amacher's compositions and installations are part of a "perceptual geography" (Amacher, 1994). She describes this concept as "seek[ing] ways of composing in which tones, originating within human anatomy, exist in their own right by becoming perceptually more than an accident of acoustic tones in the room, resulting in a conscious interplay between them. […] The object of perceptual geography is to learn to compose spatial dimensions in music— the kind of aural dimension we experience in life, but usually not in music. The idea is to create a world for the audience to enter where architecture magnifies the expressive dimensions of the music. In this approach, the phrases of the music are choreographed to sound at specific heights and locations. Tactile in presence, they appear both larger than life and small enough to touch, heard as though miles away, or felt inside the listener" (Amacher, 1994). The spatial aspect Amacher described is a unique quality that auditory distortion products offer, and the concept will be discussed in further detail in Chapter 5.

In addition to Amacher's seminal installations such as *Music for Sound Joined Rooms/Synaptic Island* (1980)[1] and recordings *Sound Characters (Making the Third Ear)* (1999), a number of recent composers and sound artists employ auditory distortion products in their work. Examples of pieces include Clarence Barlow's *…Until… Version 8 for Piccolo Flute* (1981), Jacob Kirkegaard's *Labyrinthitis* (2007), Florian Hecker's *2x3 Kanal* (2009), Christopher Haworth's *Correlation Number One* (2010), Alex Chechile's *On the Sensations of Tone I-IX* (2010-), Sergei Tcherepnin's installation *Ear Tone Box* (2013), Brian Connolly's *Invisibilia* (2014), Marcus Schmickler's *Fortuna Ribbon* (2015), and Thomas Ankersmit's *Otolit* (2015), among others. Each of these pieces evokes the phenomenon using a range of methods from acoustic instruments, analog modular

[1] The dates in this paragraph reflect the year of the artwork, not publication references.

synthesizers, and digital synthesis. While all of the pieces are linked by the intentional manipulation of the peripheral (and sometimes central) auditory system, the works exhibit the wide variety of possibilities when employing the phenomenon in music and sound installation.

1.11 CONCLUSION

The topics covered in this chapter exemplify the multidisciplinary nature of auditory distortion products. While the impetus for this work originates from music, and will eventually conclude with music, many of the topics overviewed in this chapter will be discussed in greater detail along the way. The next chapter provides a mathematical model for predicting auditory distortion product frequencies.

CHAPTER 2
Modeling the Difference Tone Frequency

2.1 INTRODUCTION

In the previous chapter, the quadratic and cubic difference tones were shown to result from the multiplication of signals. Helmholtz's seminal text *On the Sensations of Tone* (1954) provided a mathematical model for the nonlinearities of the auditory system. Fletcher (1929) simplified the model as a classical power series nonlinearity. This chapter addresses the classical power series expansion of functions and provides a solution using the multinomial theorem that yields the difference tone classes from an input signal with any number of components equal or greater than two. Finally, the model is implemented in software with the output plotted in a variety of methods, along with other helpful results. It is important to note that the model provides the frequency components of auditory distortion products, but as will be addressed in Chapter 3, their corresponding amplitude does not approximate the measured amplitude of distortion products. The following text was expanded from a presentation given at the 174th Meeting of the Acoustical Society of America conference (Chechile, 2017).

2.2 POWER SERIES NONLINEARITY MODEL WITH MULTINOMIAL THEOREM

For modeling the nonlinear response that results in auditory distortion products, let us assume the output displacement of a system $y(t)$ is a sum of terms with infinitely increasing power:

$$y(t) = \sum_{n=0}^{\infty} a_n p^n = a_0 + a_1 p^1 + a_2 p^2 + a_3 p^3 + \cdots$$

with *p* representing pressure and *a* are constants (which for simplicity will be set to 1). This equation is a classical power series. If we examine the terms, the first is equal to zero, the second term is a linear term, the third term is the quadratic term, fourth is the cubic term, fifth is the quartic term, sixth is the quintic term, and so forth. This could be expressed as:

$$y(t) = 0 + y_l(t) + y_q(t) + y_c(t) + \cdots$$

with *l* representing the linear term, *q* representing the quadratic term, and *c* representing the cubic term. Since the nonlinear system we are modeling is generated from an input of frequency components, let us consider each term as a sum of *n* number of sinusoids. For example, with a two-tone input up to the cubic term (all phases are 0°):

$$y(t) = [A_1 \sin \omega_1 t + A_2 \sin \omega_2 t]^1 + [A_1 \sin \omega_1 t + A_2 \sin \omega_2 t]^2 + [A_1 \sin \omega_1 t + A_2 \sin \omega_2 t]^3$$

the first term is the linear term, the second is the quadratic term, and the third is the cubic term. The linear term simply yields the two input sinusoids. To solve the quadratic term, we will need some trigonometric identities (Fogiel,1980):

$$A^2 \sin^2 \omega t = \frac{A^2}{2} - \frac{A^2}{2} \cos 2\omega t \tag{2.1}$$

$$\sin \omega_1 t \sin \omega_2 t = \frac{1}{2}\cos(\omega_2 - \omega_1)t - \frac{1}{2}\cos(\omega_2 + \omega_1)t \tag{2.2}$$

Through substitution in the above two-tone quadratic term $y_q(t)$, we have:

$$y_q(t) = [A_1 \sin \omega_1 t + A_2 \sin \omega_2 t]^2$$

$$y_q(t) = A_1^2 \sin^2 \omega_1 t + A_2^2 \sin^2 \omega_2 t + 2A_1 A_2 \sin \omega_1 t \sin \omega_2 t$$

$$y_q(t) = \frac{A_1^2}{2} - \frac{A_1^2}{2} \cos 2\omega_1 t + \frac{A_2^2}{2} - \frac{A_2^2}{2} \cos 2\omega_2 t + 2A_1 A_2 \left[\frac{1}{2} \cos(\omega_2 - \omega_1)t - \frac{1}{2} \cos(\omega_2 + \omega_1)t\right]$$

$$y_q(t) = \frac{A_1^2}{2} - \frac{A_1^2}{2} \cos 2\omega_1 t + \frac{A_2^2}{2} - \frac{A_2^2}{2} \cos 2\omega_2 t + A_1 A_2 [\cos(\omega_2 - \omega_1)t - \cos(\omega_2 + \omega_1)t]$$

By substituting ω_1 with f_1 and ω_2 with f_2 in our resulting polynomial, we get the quadratic difference tone f_2-f_1, the quadratic summation tone f_2+f_1, along with multiples of our input frequencies $2f_1$ and $2f_2$. Using the same trigonometric identities, a three-tone input would yield the same distortion products along with f_3-f_1, f_3+f_1, f_3-f_2, f_3+f_2, and the additional multiple $2f_3$.

Employing the multinomial theorem will help simplify solving the cubic nonlinearity. The multinomial theorem is generalized as:

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{k_1 + k_{2+\cdots+} k_m = n} \frac{n!}{k_1!\, k_2! \ldots\, k_m!} x_1^{k_1} x_2^{k_2} \ldots x_m^{k_m}$$

where k_i are integers. For the cubic term with two-tones:

$$(x_1 + x_2)^3 = \sum \frac{3!}{k_1!\, k_2!} x_1^{k_1} x_2^{k_2} \tag{2.3}$$

with which we use:

$$\begin{array}{cc} k_1 & k_2 \\ 3 & 0 \\ 0 & 3 \\ 2 & 1 \\ 1 & 2 \end{array} \tag{2.4}$$

Then we apply the k_1 and k_2 values in the 2.4 matrix to the multinomial theorem in 2.3. Remembering that 0!=1 and 1!=1, using the first row of 2.4 yields:

$$\frac{3!}{3!\,0!}x_1^3 x_2^0 = 1x_1^3$$

Using the second row of 2.4 yields:

$$\frac{3!}{0!\,3!}x_1^0 x_2^3 = 1x_2^3$$

Using the third row of 2.4 yields:

$$\frac{3!}{2!\,1!}x_1^2 x_2^1 = 3x_1^2 x_2^1$$

And using the fourth row of 2.4 yields:

$$\frac{3!}{1!\,2!}x_1^1 x_2^2 = 3x_1^1 x_2^2$$

Finally, we add these four results and replace the x values with each $A\sin\omega t$ sinusoid. Using the multinomial theorem avoids solving with distributive multiplication, simplifying the process of finding the cubic nonlinearities. This method is particularly useful when solving for a signal with three or more tones. Following our two-tone cubic nonlinearity $y_c(t)$ example:

$$y_c(t) = [A_1 \sin \omega_1 t + A_2 \sin \omega_2 t]^3$$
$$y_c(t) = A_1^3 \sin^3 \omega_1 t + A_2^3 \sin^3 \omega_2 t + 3A_1^2\, A_2 \sin^2 \omega_1 t \sin \omega_2 t + 3A_1\, A_2^2 \sin \omega_1 t \sin^2 \omega_2 t$$

where each of the four terms are the four results from the multinomial theorem. We can solve for the individual distortion products using the additional trigonometric identity given in 2.5 (Fogiel, 1980), and equation 2.6 can be derived from other well-known trigonometric identities:

$$\sin^3 \omega_A t = \frac{3}{4}\sin \omega_A t - \frac{1}{4}\sin 3\omega_A t \tag{2.5}$$

$$(\sin^2 \omega_A t)(\sin \omega_B t) = \frac{1}{2}\sin \omega_B t - \frac{1}{4}[\sin(2\omega_A - \omega_B)t - \sin(2\omega_A + \omega_B)t] \tag{2.6}$$

Continuing our solution, we replace the first two terms with 2.5 and the second two terms with 2.6, but pay careful attention when substituting 2.6 in the following step. In our cubic equation above, the sin that is squared always corresponds to ω_A in 2.6, and the sin not squared always corresponds to ω_B. Specifically, in the third term of our cubic equation, the sin that is squared is ω_1 and the sin that is not squared is ω_2, so when substituting using equation 2.6, the sin that is squared (ω_A) is ω_1 and the sin that is not squared (ω_B) is ω_2. However, in the fourth term of the cubic equation, the sin that is squared is ω_2 and the sine not squared is ω_1, so when substituting using 2.6, the ω_A is ω_2 and the ω_B is ω_1:

$$y_c(t) = A_1^3\left(\frac{3}{4}\sin \omega_1 t - \frac{1}{4}\sin 3\omega_1 t\right) + A_2^3\left(\frac{3}{4}\sin \omega_2 t - \frac{1}{4}\sin 3\omega_2 t\right)$$
$$+3A_1^2A_2\left(\frac{1}{2}\sin \omega_2 t - \frac{1}{4}[\sin(2\omega_1 - \omega_2)t - \sin(2\omega_1 + \omega_2)t]\right)$$
$$+ 3A_1A_2^2\left(\frac{1}{2}\sin \omega_1 t - \frac{1}{4}[\sin(2\omega_2 - \omega_1)t - \sin(2\omega_2 + \omega_1)t]\right)$$

Once again substituting ω_1 with f_1 and ω_2 with f_2 results in the cubic difference tones $2f_1$-f_2 and $2f_2$-f_1, the cubic summation tones $2f_1+f_2$ and $2f_2+f_1$, multiples of the input frequencies $3f_1$ and $3f_2$, and the linear terms f_1 and f_2. The same trigonometric identities are used for solving a three-tone input, along with one additional identity that was derived from other well-known functions:

$$(\sin \omega_A t)(\sin \omega_B t)(\sin \omega_C t) = \frac{1}{4}[\sin(\omega_A + \omega_B - \omega_C)t + \sin(\omega_A + \omega_C - \omega_B)t$$
$$+ \sin(\omega_B + \omega_C - \omega_A)t - \sin(\omega_A + \omega_B + \omega_C)t]$$

which results in the same distortion products listed for the two-tone input, as well as the following:

$$\begin{array}{lllll} 2f_1 - f_3 & 2f_1 + f_3 & f_1 + f_2 - f_3 & 3f_3 & f_3 \\ 2f_2 - f_3 & 2f_2 + f_3 & f_1 + f_3 - f_2 & & \\ 2f_3 - f_2 & 2f_3 + f_2 & f_2 + f_3 - f_1 & & \\ 2f_3 - f_1 & 2f_3 + f_1 & f_1 + f_2 + f_3 & & \end{array}$$

2.3 SOFTWARE IMPLEMENTATION

By applying the principles discussed in this chapter to software, we can make a series of helpful calculations and visualizations. The Mathematica script provided in Appendix A provides a number of valuable results (Chechile, 2017). The first argument for the function adpC specifies the maximum order, up to cubic (1=linear, 2=quadratic, 3=cubic), and the remaining arguments are the frequencies of the signal to be calculated in ascending order, with no maximum limit. The output of the function provides the full polynomial of the nonlinearity, including all preceding orders. For example, with the input adpC[3, 698, 921, 1051], the resulting polynomial is:

$$\frac{3}{2} + \mathrm{Cos}[130\,t] + \mathrm{Cos}[223\,t] + \mathrm{Cos}[353\,t] - \frac{1}{2}\mathrm{Cos}[1396\,t] - \mathrm{Cos}[1619\,t] - \mathrm{Cos}[1749\,t] - \frac{1}{2}\mathrm{Cos}[1842\,t] - \mathrm{Cos}[1972\,t] -$$
$$\frac{1}{2}\mathrm{Cos}[2102\,t] + \frac{3}{4}\mathrm{Sin}[345\,t] + \frac{3}{4}\mathrm{Sin}[475\,t] + \frac{3}{2}\mathrm{Sin}[568\,t] + \frac{19}{4}\mathrm{Sin}[698\,t] + \frac{3}{4}\mathrm{Sin}[791\,t] + \frac{3}{2}\mathrm{Sin}[828\,t] + \frac{19}{4}\mathrm{Sin}[921\,t] +$$
$$\frac{19}{4}\mathrm{Sin}[1051\,t] + \frac{3}{4}\mathrm{Sin}[1144\,t] + \frac{3}{4}\mathrm{Sin}[1181\,t] + \frac{3}{2}\mathrm{Sin}[1274\,t] + \frac{3}{4}\mathrm{Sin}[1404\,t] - \frac{1}{4}\mathrm{Sin}[2094\,t] - \frac{3}{4}\mathrm{Sin}[2317\,t] -$$
$$\frac{3}{4}\mathrm{Sin}[2447\,t] - \frac{3}{4}\mathrm{Sin}[2540\,t] - \frac{3}{2}\mathrm{Sin}[2670\,t] - \frac{1}{4}\mathrm{Sin}[2763\,t] - \frac{3}{4}\mathrm{Sin}[2800\,t] - \frac{3}{4}\mathrm{Sin}[2893\,t] - \frac{3}{4}\mathrm{Sin}[3023\,t] - \frac{1}{4}\mathrm{Sin}[3153\,t]$$

However, the primary tone frequency relationships are needed to know which difference tone class produces which frequency component of the polynomial:

{{Cos[2 F_1], Cos[1396]}, {Cos[2 F_2], Cos[1842]}, {Cos[2 F_3], Cos[2102]}, {Cos[$F_2 - F_1$], Cos[223]},
{Cos[$F_3 - F_1$], Cos[353]}, {Cos[$F_2 + F_1$], Cos[1619]}, {Cos[$F_3 + F_1$], Cos[1749]}, {Cos[$F_3 - F_2$], Cos[130]},
{Cos[$F_3 + F_2$], Cos[1972]}, {Sin[2 $F_2 - F_1$], Sin[1144]}, {Sin[$F_2 + F_3 - F_1$], Sin[1274]},
{Sin[2 $F_3 - F_1$], Sin[1404]}, {Sin[2 $F_2 + F_1$], Sin[2540]}, {Sin[2 $F_3 + F_1$], Sin[2800]}, {Sin[2 $F_1 - F_2$], Sin[475]},
{Sin[$F_1 + F_3 - F_2$], Sin[828]}, {Sin[2 $F_3 - F_2$], Sin[1181]}, {Sin[2 $F_1 + F_2$], Sin[2317]}, {Sin[2 $F_3 + F_2$], Sin[3023]},
{Sin[2 $F_1 - F_3$], Sin[345]}, {Sin[$F_1 + F_2 - F_3$], Sin[568]}, {Sin[2 $F_2 - F_3$], Sin[791]}, {Sin[2 $F_1 + F_3$], Sin[2447]},
{Sin[$F_1 + F_2 + F_3$], Sin[2670]}, {Sin[2 $F_2 + F_3$], Sin[2893]}, {Sin[F_1], Sin[698]}, {Sin[3 F_1], Sin[2094]},
{Sin[F_2], Sin[921]}, {Sin[3 F_2], Sin[2763]}, {Sin[F_3], Sin[1051]}, {Sin[3 F_3], Sin[3153]}}

Given our particular interest in difference tones, the function also provides an array containing only the difference tones of the polynomial and their respective frequency relationships:

{{Cos[$F_2 - F_1$], Cos[223]}, {Cos[$F_3 - F_1$], Cos[353]}, {Cos[$F_3 - F_2$], Cos[130]}, {Sin[2 $F_2 - F_1$], Sin[1144]},
{Sin[$F_2 + F_3 - F_1$], Sin[1274]}, {Sin[2 $F_3 - F_1$], Sin[1404]}, {Sin[2 $F_1 - F_2$], Sin[475]}, {Sin[$F_1 + F_3 - F_2$], Sin[828]},
{Sin[2 $F_3 - F_2$], Sin[1181]}, {Sin[2 $F_1 - F_3$], Sin[345]}, {Sin[$F_1 + F_2 - F_3$], Sin[568]}, {Sin[2 $F_2 - F_3$], Sin[791]}}

As we are interested in any perceptual issues resulting from the difference tones, a third output calculates the respective equivalent rectangular bandwidth (ERB) bands of the difference tones (discussed in Chapter 3) (Glasberg & Moore, 1990):

{{$band_3$, 130}, {$band_5$, 223}, {$band_8$, 353}, {$band_8$, 345}, {$band_9$, 475}, {$band_{11}$, 568},
{$band_{13}$, 828}, {$band_{13}$, 791}, {$band_{16}$, 1144}, {$band_{16}$, 1274}, {$band_{16}$, 1181}, {$band_{17}$, 1404}}

The function also provides a few plots visualizing the resulting distortion products. In Figure 2.1, all of the distortion products are plotted by type, with the purple bars representing the linear terms, the red bars representing the quadratic distortion products, and the blue bars representing the cubic distortion products. Figure 2.2 organizes the distortion products by frequency, with all orders combined. Additionally, the width of each bar corresponds to the numerical amplitude of the respective distortion product, although these values do not represent the perceived amplitude (to be discussed in Chapter 3). As the distortion products below the primary tones are the most audible, Figure 2.3 plots only these frequencies, with the numerical amplitude given before each frequency at the top of each bar. Finally, the script produces an audio file of the input stimulus, and separate audio clips for each of the distortion products.

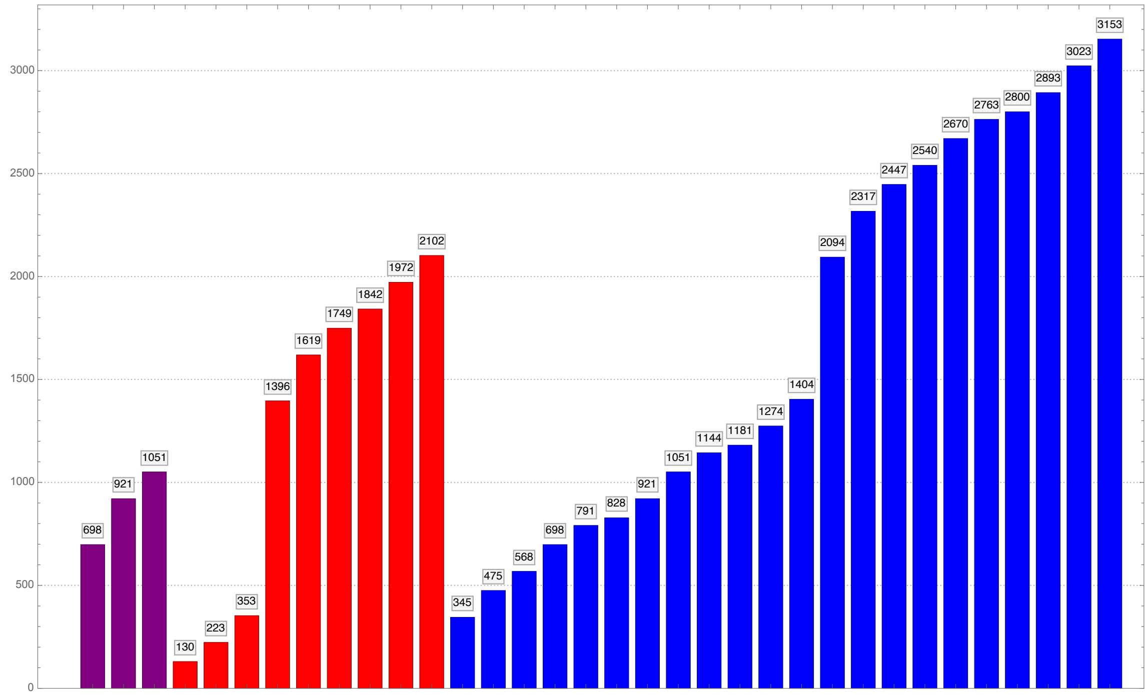


Figure 2.1. Bar plot of all distortion products produced by a three-component signal of 698 Hz, 921 Hz, and 1051 Hz (frequencies from a stimulus in the study described in Chapter 3). The purple bars represent the linear terms, the red bars represent the quadratic distortion products, and the blue bars represent the cubic distortion product, with frequencies given at the top of each bar.

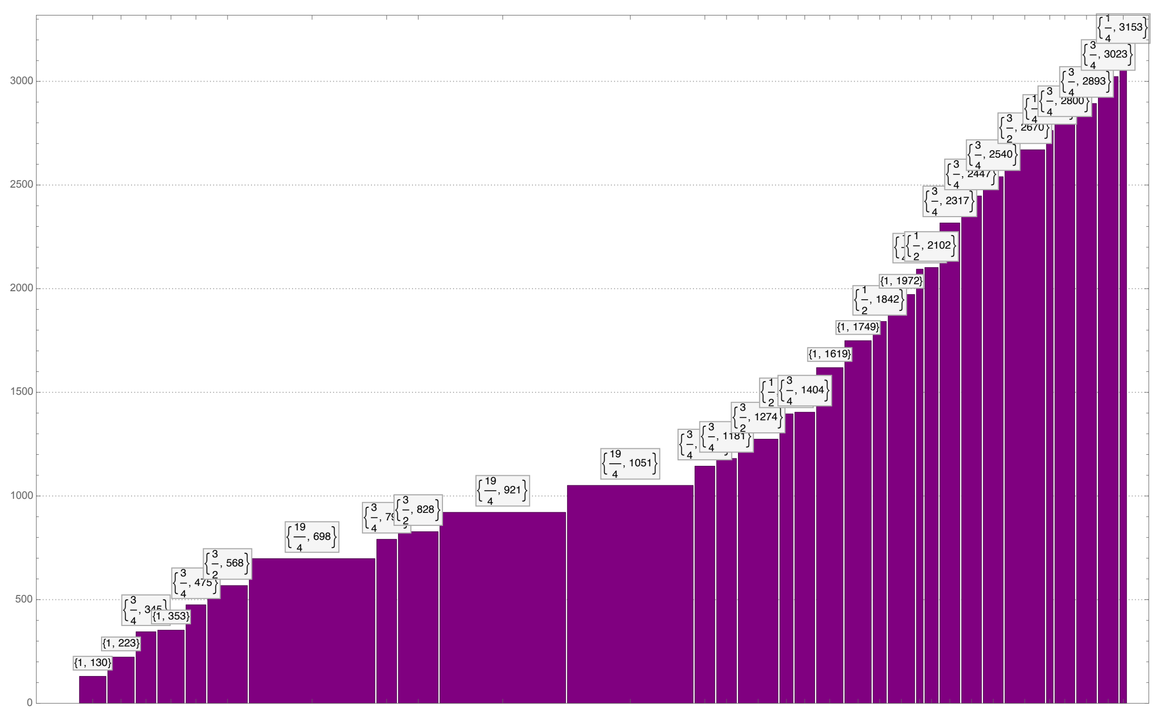


Figure 2.2. All DPs from a three-tone signal of 698 Hz, 921 Hz, and 1051 Hz, plotted by frequency. Numerical amplitude is bar width.

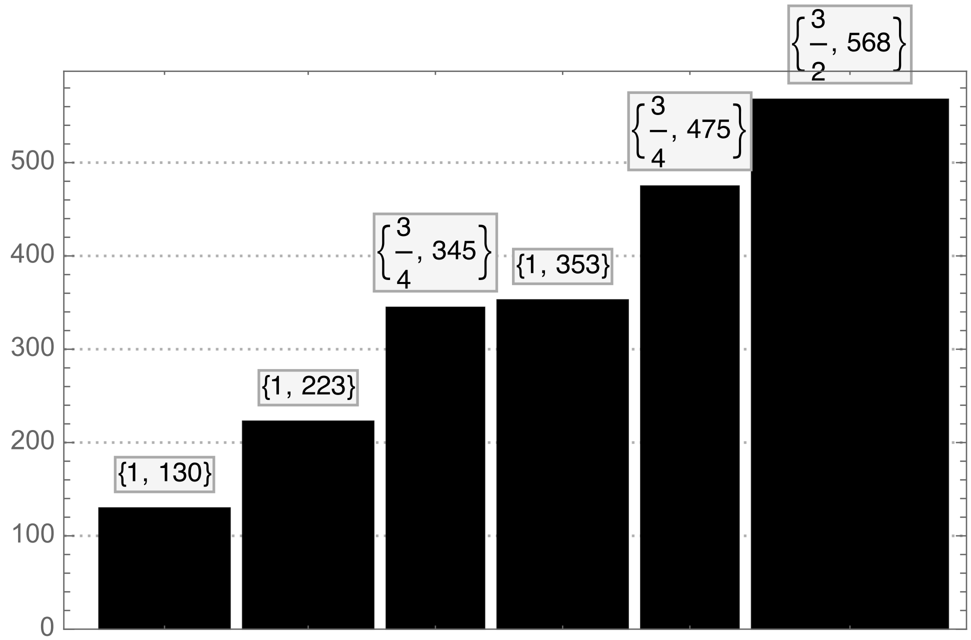


Figure 2.3. All distortion products below the lowest primary tone, from a three-tone signal consisting of 698 Hz, 921 Hz, and 1051 Hz. Numerical amplitude and frequency is provided at the top of each bar.

2.4 SUMMARY

This chapter provides a method using the multinomial theorem to find the quadratic and cubic distortion product frequencies. An implementation in Mathematica provides several additional parameters including the ERB band of the distortion products, and multiple plots of the data and subsets. The example provided in the software implementation is one of the stimulus tones used in the psychoacoustic study detailed in the next chapter.

CHAPTER 3

Perception of Difference Tones with Two and Three-Tone Stimuli

3.1 INTRODUCTION

Upon the simultaneous presentation of two or more acoustic primary tones, the ear can generate additional frequencies called difference tones as a result of auditory system nonlinearities. Given a two-tone stimulus consisting of sinusoids at f_1 and f_2 ($f_2 > f_1$), the most audible and extensively studied difference tones are at the frequencies of f_2-f_1, which is called the quadratic difference tone (QDT), and $2f_1$-f_2, which is called the cubic difference tone (CDT). Difference tones occur at other arithmetic combinations and higher orders, and their frequencies can be predicted by a classical power series nonlinearity (Helmholtz, 1863; Fletcher, 1929), however not all frequencies are audible. The odd order two-component difference tones take the general form $(n+1)f_1$-nf_2, and even order difference tones are $n(f_2$-$f_1)$ with n being a low positive integer (eg., 1, 2, 3). Generally, QDTs and CDTs will refer to f_2-f_1 and $2f_1$-f_2 respectively. Combination tones at frequencies above the lowest primary tone are far more difficult to perceive (Plomp, 1966; Goldstein, 1967; Zurek & Sachs, 1979) and therefore are not discussed in this work. Avan *et al.* (2013) reviewed the origins and current research of the phenomenon in many interconnected disciplines.

Difference tones can be measured non-invasively through quantitative psychoacoustical methods. In the method of best-beats (Wegel and Lane, 1924), a probe tone f_p is added to the primary tones at a frequency close to a combination tone and is adjusted in amplitude to produce beats. The amplitude of the probe when the beats are maximal approximates the level of the difference tone. In the cancellation method, the participant adjusts the amplitude and phase of a probe tone at the frequency of the distortion product until the tone is no longer audible (Rayleigh, 1896; Lewis and Larsen, 1937; Zwicker 1955). Like best-beats, the parameters for the cancellation tone reflect the approximate amplitude of the distortion product, but unique to cancellation is the collection of phase data. Similarly, in the cancellation of the beats method (Schouten, 1938; Goldstein, 1967), a probe tone is added a few hertz above the distortion product and the subject adjusts the amplitude and phase until a beat is heard at the number of hertz above the DP. Once a beat is established, a cancellation tone at the frequency of the distortion product is added and adjusted in amplitude and phase until the distortion product is cancelled and the beat disappears. While the techniques discussed so far all involve adding additional tones simultaneously to the primary tone complex, there are also non-simultaneous presentation techniques. In the pulsation-threshold method, the primary tones quickly alternate with a probe tone at the frequency of the distortion product (Smoorenburg, 1972). The subject adjusts the probe amplitude until the gap between the tones is no longer perceived and the distortion product frequency sounds as a continuous tone. Like the cancellation technique, the level of the probe approximates the magnitude of the distortion product, however the levels produced are typically lower than those collected by cancellation methods due to the lack of cancellation tone suppression by f_1 (Shannon and Houtgast, 1980). The pulsation-threshold method has been criticized for its assumption of linear behavior in the decay of postmasking and rise of premasking (Zwicker, 1981; Fastl and Zwicker 2007), and is not employed as frequently as cancellation techniques. Trends for non-invasively measuring distortion products moved to measuring otoacoustic emissions in the ear canal (Kemp, 1978; Kemp, 2008), and measuring far-field electrical responses in the brain from surface electrodes (Chertoff and Hecox, 1990, Rickman *et al.*, 1991; Gockel *et al.*, 2012), but psychophysical methods such as cancellation and best beats are considered ideal for they offer direct insight on the perception of the phenomenon. The methods of best-beats, cancellation, and beat-cancellation are still employed in recent work, particularly when evaluating the perception of distortion products (Pressnitzer and Patterson, 2001; Purcell *et al.*, 2007; Oxenham *et al.*, 2009).

The strength and audibility of difference tones depends on the relationship between the frequency, ratio and amplitude of the primary tones. When increasing stimulus amplitude, the even and odd-order difference tones show different slopes of growth. The classical power series nonlinearity for the QDT with equal primary levels (L_1=L_2) predicts a slope function of 2 dB per 1 dB of stimulus amplitude gain, and early studies reported slopes approximately in this range (Zwicker; 1955; Goldstein, 1967). However, a single generalization for the slope is difficult to make based on variability between subjects and measuring parameters. Using equal level primaries (L_1=L_2) and the cancellation method, Humes (1985) reported that among the five subjects measured using small f_2/f_1 ratios and a range of low stimulus amplitudes, the QDT rises at an approximate rate of 1 dB per dB of stimulus amplitude. The slope rises to 1.5 dB/dB at amplitudes ≥ 80 dB SPL. At large f_2/f_1 ratios and low stimulus amplitude levels, there is individual variability in the slope function: the QDT rises at a rate of 1dB/dB for some, and 1.6-1.7 dB/dB for others. Likewise, the slope increases at higher stimulus amplitudes. As seen in Humes' data, it is clear the slope function for QDT cancellation level and stimulus amplitude is dependent on the f_2/f_1 ratio and the general range of stimulus levels, and not as simple as the commonly reported 2dB/dB growth.

The QDT level and the primary frequency ratio f_2/f_1 are thought to be relatively independent, at least at high stimulus levels (Pickles, 2013). Goldstein (1967) measured the strength of the QDT between a ratio of 1.08 and 1.8 with f_1=2,000 Hz and found a general decrease in magnitude as the ratio increases, although there are peaks in the data as well as individual differences between the two subjects tested. In a single subject study (who was also the author), Hall (1972a) found the QDT cancellation level decreases with increasing f_2/f_1 ratios between 1.1 and 1.5 at three primary tone amplitudes (L_1=L_2=58 dB, 68 dB, and 78 dB) and four f_1 frequencies (368 Hz, 583 Hz, 926 Hz, and 1475 Hz). Hall (1972b), with an additional subject, found the QDT cancellation remains fairly even across the range of 1.42 to 1.58 in five steps. Conversely, Zwicker (1979a) measured the relationship at fixed stimulus levels of L_1=L_2=70 dB, 75 dB, 80 dB, or 85 dB with a f_1 frequency of 1620 Hz and found an approximate 10 dB increase in QDT cancellation level between the ratios ~1.12 and ~1.62 for the two subjects measured. Similarly, Humes (1985) reported an initial slight increase in cancellation level as f_2/f_1 ratios increase from

1.16 to 1.32 and then the QDT level either remains level or slightly decreases as the ratio continues to increase up to 1.68, based on median data of five subjects with f_1=1500 Hz and L_1=L_2=70 dB, 80 dB, and 90 dB. While stimulus conditions and the number of subjects vary among these studies, the overall relationship between the QDT cancellation level seems to show little dependency on the primary tone frequency ratio.

The relationship between QDT cancellation level and the f_1 frequency region is thought to be independent (Fastl & Zwicker, 2007), while experimental evidence shows both a general decrease and increase in QDT strength with a rising f_1 frequency. Measuring eight f_1 frequencies between 234 and 1180 Hz with a fixed ratio of 1.5 and stimulus level of 78 dB SPL, Hall (1972b) found the QDT decreases nearly monotonically as f_1 increases, although the data comes from one subject (the author). Hall (1972a) expanded the upper frequency range to 1475 Hz and used more fixed f_2/f_1 ratios (1.1 to 1.5) and found similar decreasing results with the f_1 frequencies 368 Hz, 583 Hz, 926 Hz, and 1475 Hz at the amplitudes L_1=L_2=78 dB, 68 dB, and 58 dB SPL. Zwicker (1979a) found a general decrease as the f_1 frequency increases between ~400 Hz and ~6,000 Hz using the parameters f_2/f_1= 1.4, L_1=80 dB, L_2=70 dB. Interestingly, within the frequency region of ~400 Hz to ~3,000 Hz, the QDT level is relatively consistent. However, in this range the QDT amplitude for both subjects increases around f_1=600 Hz and gradually decreases to a plateau until a sharp decline around 3,000 Hz. The decrease or plateauing of QDT strength with the increase of f_1 frequency contradicts findings made with non-simultaneous testing. Using a two-alternative forced choice adaptive temporal gap-masking method, Humes (1979) found the QDT amplitude increases at the middle f_1 frequency (800 Hz) before returning to its prior range (L_1=L_2=80 dB, f_1 =425 Hz, 800 Hz, 1550 Hz, f_2/f_1=1.41, four subjects).

More studies examine the odd-order CDT $2f_1$-f_2 as its behavior is far more atypical than the even-order QDT. The classical power series nonlinearity predicts CDT growth at the rate of 3 dB per 1 dB of stimulus amplitude when L_1=L_2, but psychoacoustic results show slopes that are irregular and strongly dependent on both the primary tone amplitude and ratio. As reanalyzed by Humes (1980a), early studies with L_1=L_2 show a CDT slope of 1 dB/dB (Zwicker, 1955, 1968; Goldstein, 1967), or less (Smoorenburg, 1972; Hall, 1972a). The slopes in most studies show both monotonicity and nonmonotonicity. In general, the CDT cancellation level increases with increasing amplitude of the equal amplitude primary tones, and decreases as the f_2/f_1 ratio increases, as shown in studies such as Zwicker & Harris (1990). In their report, the four subjects generally show monotonic growth of the cancellation level as a function of stimulus amplitude (L_1=L_2) between 35 dB to 80 dB, with slopes ranging 0.8 to 1.06 and increasing f_2/f_1 ratio (1.11, 1.14, 1.20, 1.27, and f_1=1620 Hz). However, studies such as Hall (1975) showed nonmonotonic growth of the CDT with stimulus amplitude (one subject, L_1=L_2, 50-70 dB SPL). The behavior with a f_2/f_1 ratio of 1.19 (f_1=840 Hz, f_2= 1,000 Hz) is nearly monotonic, but the behavior is strongly nonmonotonic at the ratios 1.25 (f_1=800 Hz, f_2= 1,000 Hz) and 1.32 (f_1=760 Hz, f_2= 1,000 Hz). In the latter cases, the CDT decreases to the lowest point around 60 dB, then increases fairly regularly until 70 dB, at which point the cancellation levels are slightly higher than the initial starting point. In the same study, Hall investigated the relationship between CDT cancellation and the f_2/f_1 ratio using a constant stimulus amplitude (L_1=L_2) and found a regular decrease in CDT cancellation tone with an increase in ratio between 1.14 and 1.47 at 56 dB and 68 dB SPL. Although, the cancellation level decreases nonmonotonically at 62 dB SPL as there is a brief increase in CDT amplitude around the ratio 1.38. Both monotonic and nonmonotonic behavior as a function of the CDT, primary tone ratio, and f_1 frequency were also observed in prior work by Hall (1972a, 1972b).

Zwicker (1981) investigated the dependence of CDT cancellation levels on primary tone amplitude and f_2/f_1 ratio across a range of low, medium, and high f_1 frequencies with up to six subjects. To summarize the equal amplitude primary tone section of Zwicker's study, the CDT level typically rises monotonically with the primary tone amplitude, remains the same or slightly rises across the f_1 frequency range measured, and decreases monotonically with increasing primary tone ratio. As reported with one subject in the low f_1 category (f_1=540 Hz), CDT levels grow monotonically with an increase of primary levels (L_1=L_2) from ~40 dB to ~80 dB SPL, with a tapering effect starting around 65 dB SPL. The CDT level decreases monotonically with an increase of primary tone f_2/f_1 ratio from 1.13 to 1.67, with the sharpest decline at the last ratio measured (L_1=L_2= 35 dB, 45 dB, 55 dB, 65 dB, 75 dB SPL). The only amplitude that produced a cancellation tone at all ratios was the maximum L_1=L_2=75 dB SPL. With a medium range frequency (f_1=1620 Hz,

L_1=L_2) in the same amplitude range, Zwicker reported monotonic CDT growth with stimulus amplitude gain, without much of the tapering effect seen at the lower f_1 range. In contrast to the single subject in the low f_1 frequency region, the data in the medium frequency range is the mean of all six subjects. The medium frequency range data reporting the relationship between CDT cancellation and primary ratio f_2/f_1 (range= 1.07 to 1.47) shows similar behavior to the low frequency region, but the cancellation levels are generally higher and Zwicker observed a gradual decline in all amplitudes after the ratio 1.11. These results are consistent with earlier work by the author (Zwicker, 1979a), which extends the ratio to 1.62 and the amplitude to a maximum 85 dB SPL. For the high frequency region (f_1= 4800 Hz), CDT cancellation tones show steep monotonic growth with increasing primary amplitude at low ratios (1.04, 1.06, 1.09), but the growth decreases at 1.12, and exhibits nonmonotonicity at the widest reported ratios of 1.16 and 1.21. Like the low f_1 frequency region, the data is reported from a single, yet different, subject. The relationship between the CDT and primary tone ratio in the high frequency region is nearly even at all primary amplitudes until the ratio of 1.10, after which the CDT levels decrease monotonically. However, the lowest two stimulus amplitudes measured (L_1=L_2, 25 dB and 35 dB SPL) show slightly nonmonotonic behavior. While not discussed in this section, the results for varied primary tone levels ($L_1 \neq L_2$) are generally nonmonotonic and exhibit arch-like behavior.

The psychoacoustic investigation of difference tones evoked with three-component signals is studied far less than the two-component variety. The focus of such studies is typically limited to the specific topics of masking/suppression (Zwicker, 1979a, 1979b), or the role of combination tones in phase effects and residue pitch (Buunen *et al.*, 1974; Buunen & Bilsen, 1974; Buunen, 1975; Pressnitzer & Patterson, 2001, Oxenham *et al.*, 2009). The latter category usually employs harmonically related primary tones. As three-component signals produce a number of other distortion products, Weber and Mellert (1975) investigated the levels of the cubic difference tone class f_1+f_2-f_3, and found similar dependence on amplitude as the two-tone CDT (f_1=950 Hz, f_2=1050 Hz, f_3=1200 Hz, L_1=L_2=L_3, four subjects). By visually extrapolating the values of a plot for one subject showing the f_1+f_2-f_3 difference tone cancellation levels as a function of stimulus amplitude (40-90 dB SPL), a reanalysis of the data demonstrates a regression slope line of 0.47 (r^2 =

0.98). The shape of the plot exhibits primarily monotonic growth with a slightly descending plateau at the two highest stimulus amplitudes, 85 dB and 90 dB SPL. The authors did not find a dependence between the $f_1+f_2-f_3$ cancellation level and the variation of the f_1 and f_2 using a fixed f_3 of 1300 Hz and a fixed center frequency between f_1 and f_2 of 1000 Hz ($L_1=L_2=L_3$, 70 dB SPL), except at the amplitude of 75 dB SPL. With fixed low and middle frequencies and a varying f_3 (f_1=930 Hz, f_2=1070 Hz, f_3=1300-1650 Hz), all four subjects demonstrate a nonmonotonic decrease of the $f_1+f_2-f_3$ cancellation level as f_3 increased. The observed behavior of the three-component stimulus is in agreement with the two-tone CDT levels.

Given the difficulty and complexity of measuring difference tones with psychoacoustic methods and the comparative ease of recording distortion product otoacoustic emissions (DPOAEs), more studies investigating three or more component signals utilize the later method (Brown & Kemp, 1984; Harris *et al.*, 1992; Cianfrone *et al.*, 1994; Kummer *et al.*, 1995; Fahey *et al.*, 2000; Martin *et al.*, 2003; Marquardt *et al.*, 2007; Bian & Scherrer, 2007; Meenderink & van der Heijden, 2010, 2011; Nuttall *et al.* 2018). When the two methods are compared, the levels for perceptual cancellation rarely coincide with the cancellation of the DPOAE (Wilson, 1980; Furst, *et al.,* 1988, Zwicker & Harris, 1990), and it is suggested the results of the two methods are not equal or proportional (Sisto *et al.*, 2018). Zwicker and Harris (1990) found the typical monotonic CDT amplitude slopes that occur with equal level primaries using psychoacoustic cancellation, but when employing DPOAE cancellation, they observed slopes that are generally nonmonotonic and lower in level. As the purpose of DPOAE studies is typically to gain a better understanding of auditory biomechanics, and the purpose of the present work is centered on the perception of difference tones, emphasis is placed on psychoacoustic methods in the present study.

The work overviewed in this section typically focuses on a limited frequency range and primary tone ratios, such as measuring CDT and/or QDT levels evoked from a small number of fixed f_1 frequencies, a few f_2 frequencies, and various stimulus amplitudes. Additionally, the psychoacoustic studies involve a small number of subjects, which often include the authors. While a wider range of conditions is desirable, concatenating past studies is problematic due to the different measuring conditions, methods, and individual differences between subjects. By focusing on detection rather than measurement, it is possible to test more than one or two types of difference tones in a single experimental session. Plomp (1965, 1966) reviewed legacy work that reported both low and high-order combination tone detection across a wide range of frequencies and ratios (Hallstrom, 1832; Konig, 1876; Meyer, 1896; Krueger, 1900; Stumpf, 1910). In these studies, the investigators used acoustic instruments to match the frequency of many combination tones. Using updated methods and technology, Plomp (1965, 1966) examined the detection of multiple orders of two-tone combination tones in both limited and wide-frequency range and ratios.

A comprehensive study on the detection of difference tones produced by two and three-tone stimuli at a fixed level sufficiently above threshold does not yet exist. By focusing on detection at a single stimulus amplitude, it is possible to learn how many classes of difference tones are perceivable by two and three-tone stimuli across a wide range of f_1 frequencies and primary tone ratios. The collected data will form a large database of usable primary tones that are known to reliably evoke multiple classes of difference tones. The database can aid the development of future perceptual research, and it can be applied to music instrument design, music composition, and sound art. The aim of the present study is to investigate the perception of specific quadratic and cubic difference tones produced by primary tones (f_1, f_2, f_3) consisting of two and three-tone stimuli across a wide frequency range, and the primary tone ratios within an octave. Of particular interest are the primary tone ratio boundaries for the detection of specific difference tones as the stimuli vary across the frequency range.

3.2 METHODS

Participants

The study consisted of 30 human subjects of ages ranging from 18-35 years (mean = 26.1 years, median = 27 years, standard deviation = 4.44 years), which were randomly split evenly into two groups of 15 subjects. Subjects volunteered in response to advertising and received a gift card for their participation in the amount of $20 per hour of involvement.

The participants were not familiar with the experiment, task, or procedure prior to volunteering, but a small minority were aware that it involved difference tones in some form. In an initial survey, subjects reported their age, level of musicianship (27 musicians), and if they had perfect pitch (four subjects reported perfect pitch to varying degrees). Subjects also reported if they were experiencing constant tinnitus, ear pain, ear infections, general hearing problems, or if they were taking any prescriptions known to change hearing. All participating subjects had normal hearing, as determined by air and bone conduction threshold measurements with a Grason Stadler GSI-61 audiometer. For inclusion, hearing thresholds were below 25 dB HL across the frequency range of 250 Hz to 8,000 Hz for air conduction, and 250 Hz to 4,000 Hz for bone conduction. An additional nine subjects participated but were rejected due to hearing thresholds exceeding the inclusion guidelines (seven subjects), not passing the training portion of the study (one subject), or an inability to focus on the assigned task (one subject). The results of these additional nine subjects were not counted in the total reported 30 participants, or reported survey responses.

Design, Procedure and Equipment

The experimental design was a detection task based on a variation of the method of best beats. To test the detection of a large number of difference tone types across a wide frequency spectrum and ratio, time efficiency was vital for the number of trials in the study. Because the design entailed measuring distortion product detection and not amplitude, the adjustable probe or cancellation tone was unnecessary. Adjusting the probe tone level, or cancellation tone amplitude/phase would have been impractical due to the time required for cancelling the beat, cancelling the distortion product, or finding the strongest beat. Instead, two fixed probe tone levels were employed in two respective groups. The fixed probe tone levels were based on reported values in past work (Warren & Egan, 1951; Hall, 1972a; Hall, 1975, Purcell *et al.*, 2007; Oxenham *et al.*, 2009). Details on the groups and probe levels will be covered in the Stimulus section.

In the design, fifty percent of the stimuli consisted of two or three-frequency primary tones with an additional probe tone at the frequency of 3 Hz above an individual distortion product. As in signal detection theory, these cases are identified as target-present cases. The remaining fifty percent of the stimuli were two or three-frequency primary tones without a probe tone, known as target-absent cases. If an audible distortion product was present, the added probe tone would produce a 3 Hz beat, which the participant was trained to detect (details to follow). This method was designed primarily for the quick detection of a large number of difference tones at a level well above threshold.

The study took two hours or less to complete for each subject and consisted of a twenty-minute hearing evaluation (the previously described survey, and air/bone conduction threshold measurements), a ten-minute training session, and an approximately hour-long experiment. During the study, the subject could take up to three five-minute breaks at will. If the subject's hearing thresholds were within a normal range, the participant engaged in a three-phase training session as described in the following paragraphs. Upon successful completion of the training, the subject proceeded to the full experiment.

Using an Apple MacBook Pro laptop running OS X 10.11.6, the first and second phases of training used custom applications written in Max/MSP 7.3.5. The third phase of training, and the main experiment used PsychoPy 3.0.0b7. The stimulus components were mixed to a single channel and were presented binaurally using RadioEar DD450 circumaural headphones at 85 dB SPL. The OS X system extension SoundFlower routed the system audio to a system software component Ableton Live 8. In Ableton Live, the signal was split into two mono tracks using a custom Max4Live plugin. Next, both channels were attenuated by -20 dB, and each channel was separately processed by an AUGraphicEQ graphic equalizer plugin.[2] The equalized audio in Ableton Live was output from an external MOTU Ultralite AVB sound card with RadioEar DD450 headphones. The total harmonic distortion from the RadioEar headphones across the 125-6300 Hz range was measured to be less than or equal to 0.6% on the left channel, and less than or equal to 0.7% on the right channel. The training and full experiment were conducted in an IAC Controlled Acoustical Environments sound attenuation booth.

In the first phase of training, the participant was instructed to note the difference between sound examples of two sinusoidal frequencies beating at 3 Hz and a single non-beating sinusoid. The frequencies were chosen at random between 500 Hz and 2000 Hz,

[2] The equalizer bands between 125 Hz to 6300 Hz were previously calibrated to 85 dB SPL.

and are clearly identified as beating or non-beating tones. In the beating examples, the participant was instructed to listen for six pulsations in the two-second sound sample, which indicated a 3 Hz beat. The subject then completed a ten-question test on the identification of beating and non-beating sounds, with a passing score equal or greater than 90%. If the participant failed, the labeled beating and non-beating examples were presented again, and the subject was given a final chance to pass the test. Any subject who failed the second attempt was dismissed as they would likely have difficulty identifying beats in the full experiment.

The second phase of training was designed for the subject to experience the sensation of a probe tone beating with a distortion product at a range of possible levels. A two-tone complex of primaries at 698 Hz and 803 Hz continuously sounded, evoking a f_2-f_1 difference tone at 105 Hz. With a volume slider, the subject adjusted the amplitude of a probe tone at 108 Hz (3 Hz higher than the f_2-f_1 QDT). No beating was present when the slider was in the minimum position, but noticeable beating would appear as the probe tone interacted with the difference tone. The subject was instructed to first slowly raise the slider until they heard a beat, then lower it until the beat disappeared, and finally to raise it again to find the approximate minimum threshold of the beat. The position of the slider was noted by the administrator.

The third and final phase of training acclimated the subject to the experimental task, input control, and stimuli. Using PsychoPy, the participant was instructed to press the right arrow on the keyboard labeled “beat” after hearing each of the five examples of beating stimuli. No other key yielded a response, so to progress the participant was required to press the correct key. Next, the same process repeated for non-beating tones, but the subject could only press the left arrow labeled “non-beat.” The five target-present beating examples were determined based on high discrimination in a pilot study and represented no more than one of a single class of distortion product (. The five target-absent non-beat examples were of the same primary tone configurations but did not contain the probe tone. The sound examples were a small subset of stimuli from the full study, were presented in random order within each respective section, and are listed in Table 3.1. After the non-beat examples ended, the third training phase was repeated once for emphasis and clarity. At the end, the subject was asked if the task was understood, and if the beating and non-beating examples were distinguishable. The full experiment began with a positive answer, or the subject was dismissed upon a negative response.

Target-Present	**Target-Absent**
519, 603, 803 (203)	519, 603, 803
698, 803 (108)	698, 803
921, 1051, 1358 (747)	921, 1051, 1358
1358, 1539, 2489 (1134)	1358, 1539, 2489
2212, 2489, 2798 (1906)	2212, 2489, 2798

Table 3.1. The five stimuli used in the practice training phase 3, with frequency in Hz. Target-present stimuli have the probe tone f_p frequency in parenthesis, and target-absent cases do not have a probe tone. The stimuli are presented in random order within each category (target-present, and target-absent, but are not mixed). The stimuli are a subset of the stimuli from the full study, and were selected due to high detectability in a pilot study.

In the PsychoPy-driven trials, the subject was instructed to listen to a sound sample and to answer whether the sound contained a beat or a non-beat. Subjects were asked to vote "beat" if they heard approximately six pulsations in the sound, although counting was not necessary. All beat and non-beat sounds were combined and presented in a unique random order each time the study was administered. The subject could only enter their response after the sound finished playing, and the response would trigger the next trial to begin. Testing of the subjects alternated between the two groups, which only differed in sound level of the probe tone. The two groups are described in detail in the following section.

Stimulus

The equivalent rectangular bandwidth (ERB) is a mathematical method for approximating the human auditory filter bandwidths. The stimuli used in the current study were built from pure tones at ERB center frequencies within the range of 500 Hz to 3,000 Hz. The ERB center frequencies were calculated by way of the number of ERBs equation

$$\text{Number of ERBs} = 21.4\ \log_{10}(0.00437F+1) \qquad (3.1)$$

with F=center frequency in Hz (Glasberg & Moore, 1990) for the respective lowest (500 Hz) and highest (3,000 Hz) frequencies. The scale was constructed by taking the ceiling of the lowest ERB number (11) and the floor for the highest ERB number (24), and performing the inverse of the number of ERBs equation

$$(10^{(E/21.4)}-1)/0.00437 \tag{3.2}$$

with E=ERB number, for every whole integer between the range (11 – 24), which resulted in the following frequencies (in hertz):

{519, 603, 698, 803, 921, 1051, 1196, 1358, 1539, 1739, 1963, 2212, 2489, 2798}

The ERB scale was first arranged into every possible combination for two and three-frequency groups, and then filtered to meet a variety of conditions. Specifically, the frequency combinations were refined to omit two or more of the same center frequencies in any given combination, and were arranged so $f_1 < f_2 < f_3$. Furthermore, the ratios of the primary frequencies ($f_2/f_1, f_3/f_1, f_3/f_2$) were limited to the range of 1.05 to 1.95. While the minimum ratio value of 1.05 was initially chosen due to its small value, the specific ERB center frequencies allowed an actual minimum ratio of 1.12. The maximum 1.95 ratio was selected because it was below the musical octave of 2.0. The result of such conditioning produced 145 primary tone stimulus combinations, consisting of 53 two-tone primary tone combinations and 92 three-tone primary tone combinations. Depending on the number of primary tones in each stimulus, the complex produced the quadratic and cubic difference tones listed in Table 3.2.a. The difference tones shown in Table 3.2.b were omitted because the difference tone frequencies were above the lowest primary tone. From all of the primary tone combinations, there were 715 difference tones from two-tone primaries and three-tone primaries investigated in this study.[3] With the additional 715 target-absent conditions (primary tones only), along with the 36 alternative probe amplitude conditions,

[3] The number was technically slightly lower at 713 unique combinations because two cases of the same primary tone combinations produced the same difference tone by different frequency relationships. Specifically, the primary tones 803 Hz, 1051 Hz, and 1358 Hz produced the difference tone 248 Hz by both f_2-f_1 and $2f_1$-f_3, and the same primary tones also produced the difference tone 555 by both f_3-f_1 and $2f_1$-f_2. The two identical stimuli were tested twice to account for each type of difference tone, therefore the grand total was 715. The identical pairs shared the same detection results in both instances.

and their respective 36 target-absent conditions, the study consisted of 1502 total trials. Each stimulus was presented once in a random order.

Included
f_2-f_1
$2f_1$-f_2
f_3-f_1
f_3-f_2
$2f_1$-f_3
f_1+f_2-f_3
$2f_2$-f_3

Excluded
$2f_2$-f_1
f_2+f_3-f_1
$2f_3$-f_1
f_1+f_3-f_2
$2f_3$-f_2

Table 3.2.a. (left) Difference tones investigated with $f_1<f_2<f_3$. Table 3.2.b. (right) Difference tones excluded because the resulting auditory distortion product was a frequency greater than the lowest primary tone, with $f_1<f_2<f_3$.

The stimuli were generated in Mathematica 11.1.1 at a 48 kHz sampling rate and 24-bit depth. Each stimulus was two seconds in duration with 20ms rise time and 20ms decay time to reduce the generation of components at frequencies outside of the primary tones. All f_1, f_2, and f_3 sinusoids began with the same starting phase 0°. The overall stimulus level was 85 dB SPL with equivalent primary tone levels (L_1=L_2=L_3), and the level of the probe tone, L_p, fixed at one of two levels L_p=L_1-20 dB for participant Group A, and L_p=L_1-30 dB for participant Group B. The probe tone and primary tone levels were calculated using the following generalized formulas:

$$nA+rA = 1 \tag{3.3}$$

$$A = \frac{1}{n+r} \tag{3.4}$$

with r representing the desired probe tone linear amplitude reduction, and n representing the number of primary tones (excluding the probe tone). For example, in a three primary tone complex with a probe tone at L_p=L_1-20 dB (r = 0.1), the primary tone amplitude was calculated by:

$$A = \frac{1}{3+0.1} = 0.3225806452 \tag{3.5}$$

and the probe tone was calculated by:

$$rA = \frac{1}{1+\frac{3}{0.1}} = 0.0322580645 \tag{3.6}$$

and all of the resulting amplitudes were multiplied by 0.2 for scaling within the dynamic range of the 24-bit amplitude resolution. Consequently, the relative primary tone values calculated with equation 3.5 would be 0.064516129 after scaling, and the relative probe stimulus values would be 0.0064516129 after scaling. To achieve equal power amplitude for the stimulus complex, the relative amplitude values of the individual primary and probe tones were multiplied by a value corresponding to peak amplitude divided by the root mean square (RMS) of the individual components (0.2/RMS). The RMS was calculated by the following equation:

$$\mathrm{RMS} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}{n}}$$

For example, in the same three-primary and L_p=L_1-20 dB probe tone case, the RMS is 0.05596563, so the relative amplitudes of the primary and probe tones were multiplied by 0.2/0.05596563 = 3.5736. In the two primary tone complex with a L_p=L_1-20 dB probe tone, both the relative primary and the probe tone values were multiplied by 0.2/0.07795576 = 2.5656. The target-absent cases were also adjusted for equal power amplitude. The two-tone primary complex was multiplied by 0.2/0.1 = 2, and the three-tone primary complex was multiplied by 0.2/0.0666667 = 3.0, which were also the same multipliers for both Group A and Group B target-absent cases. For the stimulus in Group B, in which the probe tone is L_p=L_1-30 dB, the relative three-tone and probe values were multiplied by 0.2/0.0571423162 = 3.5, and the two-tone and probe values were multiplied by 0.2/0.803988507 = 2.4876. The adjustments for equal power amplitude produced a two-tone stimulus at 84 dB SPL overall, a three-tone stimulus at 85 dB SPL overall, and a four-tone stimulus (three primaries and a probe tone) at 86 dB SPL overall. To achieve 85 dB SPL for all stimuli, the 84 dB two-tone and the 86 dB four-tone stimuli were further

f_1	f_2	f_3	f_p
519	603	698	98
603	698		98
603	698	803	596
603	803	921	321
603	921		288
698	921	1196	649
698	1051	1358	348
698	1358		41
803	921		121
803	921	1051	251
803	921	1539	739
803	1051	1539	70
803	1196	1358	396
921	1051	1196	779
921	1358	1539	184
1051	1196	1539	566
1051	1196	1963	909
1051	1358	1739	673
1051	1358	1739	691
1051	1539	1963	491
1051	1963		142
1051	1963		915
1196	1358	2212	345
1196	1358	2212	507
1196	1539	1963	432
1196	1539	1963	856
1358	1539	2212	676
1358	1963	2212	252
1358	2489		1134
1539	1739	2489	1342
1539	1963	2489	427
1739	1963	2489	227
1739	1963	2798	907
1739	2212		1269
2212	2489	2798	1629
2212	2489	2798	2183

Table 3.3. Stimuli used in cross-group data, with probe frequencies at the opposite amplitude of the group in which they were presented. The thirty-six frequency combinations were presented once with and once without the probe f_p, and played in random order along with the other stimuli.

multiplied by values corresponding to +1 dB and -1 dB respectively.

For facilitating a comparison in the detection rate between the two groups, an additional 36 randomly selected stimuli with the alternate probe level were added to each group, along with their respective target-absent cases. For example, in the L_p=L_1-20 dB group, there were an additional 36 stimuli with a probe tone at L_p=L_1-30 dB, and 36 additional stimuli at the same primary frequencies but without the probe tone. The random selection of the cross-group stimuli was performed by a custom program built in Max/MSP, and the results are listed in Table 3.3.

Data Collection and Analysis

The dependent variable measure was the subject response regarding the detection of a beat (beat or no beat). The independent variables were the presence or absence of a probe stimulus (target-present or target-absent), the level of the probe tone (Group A with $L_p = L_1$-20 dB or Group B with $L_p = L_1$-30 dB), and the number of difference tones tested (715 total within each group, between nine difference tone classes). Another independent manipulation included the subgroup of 36 trials within each group with a probe tone at the non-dominant level (e.g., the 36 trials with $L_p = L_1$-30 dB in Group A, which primarily has a probe of $L_p = L_1$-20 dB), and their respective target-absent cases.

In the target-present condition, the hit rate of a particular difference tone trial was determined by calculating the mean of all hits across the 15 cases. The miss rate was calculated by subtracting the hit rate from one. Before calculating the false alarm rate and correct rejection rate, the sum of all cases of target-absent trials and the number of cases of a particular target-absent trial were found. The number of target-absent trials varied from two to nine instances. To find the sum of cases, the number of correct identifications of a particular non-beat trial across all subjects and the number of times that stimulus tone combination was presented were added. The maximum total for each particular target-absent stimulus was found by multiplying the number of times that stimulus was presented by 15. The false alarm rate was calculated by subtracting the sum of cases from the maximum total and divided the difference by the maximum total. The correct rejection rate was calculated by dividing the sum of cases by the max total, which is equivalent to one minus the false alarm rate. Finally, after calculating the hit rate and miss rate from the target present cases and the false alarm rate and correct rejection rate from the target-absent cases, the discrimination index (to be discussed in the following paragraph) was calculated by subtracting the false alarm rate from the hit rate.

The collected data was analyzed from a Bayesian perspective, with a Bayes factor used for statistical decisions regarding the detection of beats. Through calculating a minimum threshold level, θ_{min}, one can find a discrimination index that is statistically reliable (i.e., highly probable of detection in general based on the sample of test trials). If the estimated discrimination index is equal or larger than the minimum threshold, it has a Bayes factor that is greater than 19. The true population discrimination index δ is a difference between the population of hit proportions ϕ_{hit} and the population of false alarm proportions ϕ_{FA}, also known as the hit rate and false alarm rate proportions for the true population. A population discrimination index ≤ 0 corresponds to the case of the null hypothesis H_0 for that condition. A discrimination index > 0 corresponds to the alternative hypothesis H_1 for the detection of a beat. The prior probabilities for the null and alternative hypotheses are assumed to be equal, i.e., $P(H_0) = P(\mathrm{H}_1) = 0.5$. A sample of target-present and a separate sample of target-absent trials allows posterior probabilities, denoted as $P(H_0 \mid D)$ and $P(H_1 \mid D)$ where D denotes the data. The posterior odds ratio between hypotheses after obtaining data is

$$\frac{P(H_1 \mid D)}{P(H_0 \mid D)} = \frac{P(H_1)}{P(H_0)} \times BF_{10} \tag{3.7}$$

where BF_{10} signifies the Bayes factor (Jeffreys, 1961; Kass & Raftery, 1995). As equation 3.7 states, the ratio of posterior odds is equivalent to the prior odds ratio times the Bayes factor. The Bayes factor is the probability ratio for the data between the null and alternative hypotheses, and is represented as

$$BF_{10} = \frac{P(D \mid H_1)}{P(D \mid H_0)} \tag{3.8}$$

Both the null and alternative hypotheses are credible options from a Bayesian perspective. A large Bayes factor yields the selection of the alternative hypothesis. The null hypothesis is selected if the inverse of the Bayes factor is large:

$$BF_{01} = \frac{1}{BF_{10}} = \frac{P(D \mid H_0)}{P(D \mid H_1)} \tag{3.9}$$

The magnitude of the Bayes factor aids the decision regarding which hypothesis to select, with strong evidence corresponding to a level of 19 or greater (Jeffreys, 1961).

In our analysis, the prior odds ratio is

$$\frac{P(H_1)}{P(H_0)} = \frac{0.5}{0.5} = 1 \tag{3.10}$$

as the null and alternative hypotheses are both 0.5. Therefore, equation 3.7 yields a Bayes factor equal to the posterior odds ratio. Also, as the null and alternative hypotheses are mutually exclusive and exhaustive, the posterior probability for the alternative hypothesis is equal to one minus the posterior probability for the null hypothesis:

$$P(H_1 \mid D) = 1\text{-}P(H_0 \mid \mathrm{D}) \tag{3.11}$$

The probability is high that the alternative hypothesis is true if the $BF_{10} \geq 19$. A Bayes factor greater than 19 corresponds to cases with the posterior probability that the H_1 is 0.95 or greater is substantiated by the following:

$$BF_{10} \geq 19$$

$$\frac{P(H_1 \mid D)}{1-P(H_1 \mid D)} \geq 19$$

$$P(H_1 \mid D) \geq 19 - 19P(H_1 \mid D)$$

$$P(H_1 \mid D) \geq \frac{19}{20} = 0.95$$

$$P(\delta > 0 \mid D) \geq 0.95$$

Using the R software package, the discrimination index threshold $\theta_{\min}$ was found by performing Monte Carlo simulations. In each group, 15 sets of target-present conditions were collected for every combination of primary tones, as there were no repeat trials for a single subject. Because the detection of up to seven classes of difference tones was tested per frequency combination, there were more trials of target-absent conditions without the probe tone for each primary tone combination. In each group, there were three subsets of target-absent cases. For cases with two primary tones there were 15x57=855 trials, for the specific $2f_2$-f_3 difference tone class there were 15x61=915 trials, and for all other three-tone difference tone classes there were 15x96=1440 trials. The probability distribution for the population false-alarm rate ϕ_{FA} was simulated using the total number of false-alarm responses for each of these cases in each group. In each group and subset of target-absent conditions, R was used to create vectors of 50,000 random drawings from the appropriate posterior distribution for the random variable. In Bayesian statistics, the posterior distribution for the binomial trial test is a beta distribution (Raiffa & Schlaifer, 1961), so one can use the rbeta(50000, a, b) command in R to enable the random sampling (the shape coefficients for defining a specific beta distribution are *a* and *b*). To simulate the posterior distribution of the target-present cases, an additional set of random vectors was collected. As the target-present condition contained only 15 trials per group, vectors for either four, five, or six correct detections of the difference tone were created. In each case, 50,000 random values from the appropriate beta distribution were obtained. Random values for the true population discrimination index were also produced in R. For example, if one has an R vector of 50,000 random values for a specific category of false alarms (FA1), and another R vector of 50,000 random values with five correct detections and ten misses (Hit1), then one can create a resulting vector of 50,000 random values for the true population discrimination index (D1) by evaluating D1=Hit1-FA1 in R. The posterior probability for the true population discrimination index that is greater than zero can be estimated by the proportion of the 50,000 values in D1 that are positive. If the posterior probability is less than 0.95, one increases the number of correct detections by a value of one and decreases the number of misses by a value of one and creates a new set of 50,000 random values for the target present condition (Hit2). With the new Hit2 vector, one can evaluate D2=Hit2-FA1, and continue the process until finding a vector with greater than

95% of positive values. At that point, the estimated minimum threshold θ_{min} can be found by calculating the mean of the vector. In the present study, the false alarm rates were different for each of the conditions, so the thresholds varied. The resulting threshold values for Group A were 0.1693 for the two primary tone classes, 0.2065 for the $2f_2$-f_3 class, and 0.2294 for the rest of the three-tone difference tone classes. For Group B, the corresponding values were 0.1868, 0.1828, and 0.1968. If the difference tone discrimination index was above these minimum threshold values, the detection of the distortion product was statistically reliable, meaning the Bayes factor for the detection of the difference tone was greater than 19.

3.3 RESULTS

Let us first examine the results from a broad perspective before narrowing focus. Table 3.4 shows the mean discrimination indices with the standard error of means for each distortion product class investigated, and the mean discrimination indices for each group. In Group A, the three-tone difference tones with mean discrimination indices equal to or greater than the minimum threshold are $2f_1$-f_2, f_1+f_2-f_3, f_3-f_1, f_3-f_2, and both two-tone difference tones $2f_1$-f_2 and f_2-f_1. Therefore, six of the nine types of difference tones in Group A have mean discrimination indices that are statistically reliable. In Group B, the mean discrimination indices are equal to or greater than the minimum threshold for the three-tone difference tones $2f_1$-f_2, f_1+f_2-f_3, f_2-f_1, f_3-f_1, f_3-f_2, and both two-tone difference tones $2f_1$-f_2 and f_2-f_1. That said, the difference tone classes with a mean discrimination index below threshold still have individual cases that are statistically reliable, therefore they offer useful information for detection.

DP	Group A	Group B
$2f_1$-f_2 (2t)	0.4397 (0.0468)	0.4017 (0.0296)
f_2-f_1 (2t)	0.4548 (0.0271)	0.5715 (0.0330)
$2f_1$-f_2	0.4127 (0.0254)	0.2615 (0.0182)
$2f_1$-f_3	0.1011 (0.0158)	0.1253 (0.0152)
$2f_2$-f_3	0.1726 (0.0211)	0.1188 (0.0194)
f_1+f_2-f_3	0.3141 (0.0216)	0.2949 (0.0227)
f_2-f_1	0.2199 (0.0164)	0.4586 (0.0257)
f_3-f_1	0.2438 (0.0173)	0.2680 (0.0210)
f_3-f_2	0.3504 (0.0207)	0.5564 (0.0220)
Total	0.2913 (0.0087)	0.3344 (0.0095)

Table 3.4. Mean discrimination indices for distortion products in Group A and Group B with standard error of mean values in parenthesis. All stimuli are 3-tone unless labeled 2t.

Tables 3.5-3.8 provide the discrimination indices for each combination of primary tones and their respective f_2/f_1 and f_3/f_1 ratios. The last column in each table shows the total number of distortion products above threshold for a particular primary tone combination, with a maximum value of two in the two-tone cases, and a maximum value of seven in the three-tone conditions. The tables are color coded with the intensity of red representing the ratio value, as well as the number of difference tone classes above threshold for a given stimulus. If a discrimination index is not present, that particular difference tone was not included in the study. Overall, there are trends for detection as a function of primary tone ratio, with non-detection often occurring at larger ratios. Also, it should be noted that the number of ratios tested naturally decreases in the higher f_1 primary tone blocks due to the hard ceiling effect around 3kHz. The following paragraphs examine the ratio ranges above threshold for each of the nine difference tone classes across the 12 or 13 f_1 primary frequencies. The range is found by inspecting ratio boundaries between clusters of detection. For example, if all difference tones with the same f_1 frequency are above threshold except the first and last primary tone combinations, the ratio range is determined by the sequential primaries above threshold between the cases of non-detection.

f1	f2	2f1-f2	D.I.%	f2/f1	f2-f1	D.I.%	f2/f1	Total
519	603	435	0.7333	1.162	84	0.5333	1.162	2
519	698	340	0.7	1.345	179	0.3667	1.345	2
519	803	235	0.1667		284	0.5667	1.547	1
519	921	117	0.2	1.775	402	0.1333		1
603	698	508	0.7111	1.158	95	0.5778	1.158	2
603	803	403	0.2	1.332	200	0.2	1.332	2
603	921	285	0.3333	1.527	318	0.6	1.527	2
603	1051	155	0.2333	1.743	448	0.2333	1.743	2
698	803	593	0.7667	1.15	105	0.7667	1.15	2
698	921	475	0.7333	1.319	223	0.7333	1.319	2
698	1051	345	0.2667	1.506	353	0.4	1.506	2
698	1196	200	0.0333		498	0.2333	1.713	1
698	1358	38	0.1333		660	0.2667	1.946	1
803	921	685	0.8444	1.147	118	0.5111	1.147	2
803	1051	555	0.7667	1.309	248	0.4333	1.309	2
803	1196	410	0.1667		393	0.6333	1.489	1
803	1358	248	0.3333	1.691	555	0.8	1.691	2
803	1539	67	0.0667		736	0.1333		0
921	1051	791	0.8	1.141	130	0.6	1.141	2
921	1196	646	0.9	1.299	275	0.6333	1.299	2
921	1358	484	0.4667	1.474	437	0.2667	1.474	2
921	1539	303	0.2333	1.671	618	0.5	1.671	2
921	1739	103	0.1667		818	0.1		0
1051	1196	906	0.8667	1.138	145	0.5333	1.138	2
1051	1358	744	0.8	1.292	307	0.4667	1.292	2
1051	1539	563	0.2	1.464	488	0.4	1.464	2
1051	1739	363	0.1667		688	0.5	1.655	1
1051	1963	139	0.0833		912	0.2167	1.868	1
1196	1358	1034	0.9333	1.135	162	0.7333	1.135	2
1196	1539	853	0.8333	1.287	343	0.5	1.287	2
1196	1739	653	0		543	0.1333		0
1196	1963	429	0.1667		767	0.4333	1.641	1
1196	2212	180	0.0333		1016	0.1667		0
1358	1539	1177	0.8667	1.133	181	0.7333	1.133	2
1358	1739	977	0.6333	1.281	381	0.3	1.281	2
1358	1963	753	0.1		605	0.1667		0
1358	2212	504	0		854	0.4	1.629	1
1358	2489	227	-0.0444		1131	0.3556	1.833	1
1539	1739	1339	0.8333	1.13	200	0.5667	1.13	2
1539	1963	1115	0.5667	1.276	424	0.3667	1.276	2
1539	2212	866	0.2333	1.437	673	0.5	1.437	2
1539	2489	589	-0.0333		950	0.4333	1.617	1
1539	2798	280	0.0333		1259	0.3667	1.818	1
1739	1963	1515	0.9667	1.129	224	0.7	1.129	2
1739	2212	1266	0.6444	1.272	473	0.6444	1.272	2
1739	2489	989	0.3333	1.431	750	0.2	1.431	2
1739	2798	680	-0.0667		1059	0.4	1.609	1
1963	2212	1714	0.9333	1.127	249	0.6667	1.127	2
1963	2489	1437	0.4667	1.268	526	0.5333	1.268	2
1963	2798	1128	0.2667	1.425	835	0.3333	1.425	2
2212	2489	1935	0.8	1.125	277	0.6	1.125	2
2212	2798	1626	0.8	1.265	586	0.8	1.265	2
2489	2798	2180	0.9333	1.124	309	0.7333	1.124	2
Min:				1.124			1.124	
Max:				1.775			1.946	

Table 3.5. Two-tone difference tones from Group A showing the discrimination indices for the $2f_1$-f_2 CDT and f_2-f_1 QDT classes. The primary tone ratios are only displayed if the given difference tone is above threshold, and accordingly, the blank ratios represent difference tones below threshold. The rightmost column shows the total number of difference tones above threshold for the row's set of primary tones (maximum value is 2). The minimum and maximum ratios in each class are in the final rows.

f1	f2	f3	2f1-f2	D.I.%	f2/f1	f3/f1	2f1-f3	D.I.%	f2/f1	f3/f1	2f2-f3	D.I.%	f2/f1	f3/f1	f1+f2-f3	D.I.%	f2/f1	f3/f1	f2-f1	D.I.%	f2/f1	f3/f1	f3-f1	D.I.%	f2/f1	f3/f1	f3-f2	D.I.%	f2/f1	f3/f1	Total
519	603	698	435	0.3417	1.162	1.345	340	0.2083			508	0.075			424	0.2083			84	0.1417			179	0.2083			95	0.4083	1.162	1.345	2
519	603	803	435	0.6095	1.162	1.547	235	0.0762			403	0.2762	1.162	1.547	319	0.5429	1.162	1.547	84	0.1429			284	0.4762	1.162	1.547	200	0.2095			4
519	603	921	435	0.5714	1.162	1.775	117	0.2381	1.162	1.775	285	0.2381	1.162	1.775	201	0.0381			84	0.2381	1.162	1.775	402	0.0381			318	0.3048	1.162	1.775	5
519	698	803	340	0.2778	1.345	1.547	235	0.1444			593				414	0.2778	1.345	1.547	179	0.1444			284	0.2111			105	0.2111			2
519	698	921	340	0.4381	1.345	1.775	117	0.1048			475	-0.0952			296	0.5714	1.345	1.775	179	0.1714			402	0.2381	1.345	1.775	223	0.3714	1.345	1.775	4
519	803	921	235	0.1111			117	0.1111			685				401	0.3778	1.547	1.775	284	0.1778			402	0.0444			118	0.3778	1.547	1.775	2
603	698	803	508	0.2667	1.158	1.332	403	0.2			593	0			498	0.2			95	0.2667	1.158	1.332	200	0.1333			105	0.3333	1.158	1.332	3
603	698	921	508	0.4952	1.158	1.527	285	0.1619			475	0.3619	1.158	1.527	380	0.5619	1.158	1.527	95	0.2286			318	0.2952	1.158	1.527	223	0.6286	1.158	1.527	5
603	698	1051	508	0.4476	1.158	1.743	155	-0.0857			345	0.2476	1.158	1.743	250	-0.019			95	0.1143			448	0.181			353	0.1143			2
603	803	921	403	0.5619	1.332	1.527	285	-0.0381			685				485	0.5619	1.332	1.527	200	0.2952	1.332	1.527	318	0.3619	1.332	1.527	118	0.4286	1.332	1.527	5
603	803	1051	403	0.4952	1.332	1.743	155	0.0952			555	0.0952			355	0.4286	1.332	1.743	200	0.0952			448	0.2286			248	0.5619	1.332	1.743	3
603	921	1051	285	0.2333	1.527	1.743	155	0.1			791				473	0.5667	1.527	1.743	318	0.1			448	0.1			130	0.4333	1.527	1.743	3
698	803	921	593	0.2762	1.15	1.319	475	0.2095			685	0.0095			580	0.2762	1.15	1.319	105	0.1429			223	0.2095			118	0.0095			2
698	803	1051	593	0.4095	1.15	1.506	345	0.1429			555	0.1429			450	0.5429	1.15	1.506	105	0.2762	1.15	1.506	353	0.1429			248	0.5429	1.15	1.506	4
698	803	1196	593	0.2571	1.15	1.713	200	-0.0762			410	0.3238	1.15	1.713	305	0.1238			105	0.3238	1.15	1.713	498	0.0571			393	0.3238	1.15	1.713	4
698	803	1358	593	0.6667	1.15	1.946	38	0.0667			248	0			143	0			105	0.4667	1.15	1.946	660	0.1333			555	0.0667			2
698	921	1051	475	0.5444	1.319	1.506	345	0.3444	1.319	1.506	791				568	0.4111	1.319	1.506	223	0.1444			353	0.3444	1.319	1.506	130	0.2778	1.319	1.506	5
698	921	1196	475	0.475	1.319	1.713	200	0.075			646	0.1417			423	0.5417	1.319	1.713	223	0.4083	1.319	1.713	498	0.275	1.319	1.713	275	0.6083	1.319	1.713	5
698	921	1358	475	0.3333	1.319	1.946	38	0			484	0.2			261	-0.0667			223	0.2			660	0			437	0.2			1
698	1051	1196	345	0.1667			200	-0.0333			906				553	0.2333	1.506	1.713	353	-0.1667			498	0.2333	1.506	1.713	145	0.3667	1.506	1.713	3
698	1051	1358	345	-0.0571			38	-0.1905			744				391	0.2095			353	0.0095			660	0.0762			307	0.2762	1.506	1.946	1
698	1196	1358	200	0.0667			38	-0.0667			1034				536	0.2			498	-0.0667			660	-0.1333			162	0.6667	1.713	1.946	1
803	921	1051	685	0.3583	1.147	1.309	555	0.3583	1.147	1.309	791	0.0917			673	0.425	1.147	1.309	118	0.225			248	0.1583			130	0.025			3
803	921	1196	685	0.4381	1.147	1.489	410	0.1714			646	0.0381			528	0.4381	1.147	1.489	118	0.1048			393	0.1714			275	0.3048	1.147	1.489	3
803	921	1358	685	0.6857	1.147	1.691	248	0.019			484	0.1524			366	0.419	1.147	1.691	118	0.219			555	0.219			437	0.219			2
803	921	1539	685	0.7083	1.147	1.917	67	0.1083			303	0.175			185	0.2417	1.147	1.917	118	0.175			736	0.1083			618	0.1083			2
803	1051	1196	555	0.4889	1.309	1.489	410	0.3556	1.309	1.489	906				658	0.4222	1.309	1.489	248	0.2889	1.309	1.489	393	0.2889	1.309	1.489	145	0.4889	1.309	1.489	6
803	1051	1358	555	0.5714	1.309	1.691	248	0.3714	1.309	1.691	744	0.1048			496	0.5048	1.309	1.691	248	0.2381	1.309	1.691	555	0.5048	1.309	1.691	307	0.6381	1.309	1.691	6
803	1051	1539	555	0.1167			67	0.1167			563	0.05			315	0.1167			248	0.1167			736	0.05			488	0.05			0
803	1196	1358	410	0.2476	1.489	1.691	248	0.0476			1034				641	0.4476	1.489	1.691	393	0.181			555	0.1143			162	0.5143	1.489	1.691	3
803	1196	1539	410	0.2			67	0			853				460	0.4667	1.489	1.917	393	0.1333			736	0			343	0.2667	1.489	1.917	2
803	1358	1539	248	-0.1778			67	0.0222			1177				622	0.1556			555	0.0222			736	0.0889			181	0.3556	1.691	1.917	1
921	1051	1196	791	0.3167	1.141	1.299	646	0.1833			906	0.05			776	0.25	1.141	1.299	130	0.1167			275	0.3167	1.141	1.299	145	0.1167			3
921	1051	1358	791	0.7714	1.141	1.474	484	0.0381			744	0.3714	1.141	1.474	614	0.7048	1.141	1.474	130	0.3714	1.141	1.474	437	0.3714	1.141	1.474	307	0.5048	1.141	1.474	6
921	1051	1539	791	0.5714	1.141	1.671	303	-0.0286			563	0.1048			433	0.5048	1.141	1.671	130	0.4381	1.141	1.671	618	0.2381	1.141	1.671	488	0.1714			4
921	1051	1739	791	0.6762	1.141	1.888	103	0.0762			363	0.0762			233	0.0762			130	0.0762			818	0.1429			688	0.2762	1.141	1.888	2
921	1196	1358	646	0.3556	1.299	1.474	484	0.0889			1034				759	0.4222	1.299	1.474	275	0.2222			437	0.2889	1.299	1.474	162	0.2889	1.299	1.474	4
921	1196	1539	646	0.4762	1.299	1.671	303	0.0095			853	0.2095	1.299	1.671	578	0.4762	1.299	1.671	275	0.1429			618	0.0762			343	0.4762	1.299	1.671	4
921	1196	1739	646	0.2476	1.299	1.888	103	0.1143			653	0.0476			378	-0.0857			275	-0.019			818	0.1143			543	0.1143			1
921	1358	1539	484	0.2286			303	0.2952	1.474	1.671	1177				740	0.4286	1.474	1.671	437	0.2286			618	0.3619	1.474	1.671	181	0.5619	1.474	1.671	4
921	1358	1739	484	0.2778	1.474	1.888	103	-0.0556			977				540	0.2778	1.474	1.888	437	0.2111			818	0.0111			381	0.4111	1.474	1.888	3
921	1539	1739	303	0.1111			103	0.1111			1339				721	0.4444	1.671	1.888	618	0.2444	1.671	1.888	818	0.2444	1.671	1.888	200	0.5111	1.671	1.888	4
1051	1196	1358	906	0.4857	1.138	1.292	744	0.3524	1.138	1.292	1034	0.0857			889	0.4857	1.138	1.292	145	0.019			307	0.219			162	0.0857			3
1051	1196	1539	906	0.65	1.138	1.464	563	0.25	1.138	1.464	853	0.3833	1.138	1.464	708	0.3833	1.138	1.464	145	0.45	1.138	1.464	488	0.3167	1.138	1.464	343	0.3833	1.138	1.464	7
1051	1196	1739	906	0.6952	1.138	1.655	363	-0.0381			653	0.1619			508	0.2286			145	0.4286	1.138	1.655	688	0.4286	1.138	1.655	543	0.3619	1.138	1.655	4
1051	1196	1963	906	0.45	1.138	1.868	139	-0.0833			429	-0.0833			284	-0.0833			145	0.05			912	-0.0167			767	-0.0833			1
1051	1358	1539	744	0.6222	1.292	1.464	563	0.2222			1177				870	0.6222	1.292	1.464	307	0.4222	1.292	1.464	488	0.2222			181	0.4222	1.292	1.464	4
1051	1358	1739	744	0.3704	1.292	1.655	363	0.237	1.292	1.655	977	0.1037			670	0.3037	1.292	1.655	307	0.237	1.292	1.655	688	0.3704	1.292	1.655	381	0.237	1.292	1.655	6
1051	1358	1963	744	0.381	1.292	1.868	139	0.2476	1.292	1.868	753	0.381	1.292	1.868	446	-0.1524			307	0.2476	1.292	1.868	912	0.181			605	0.1143			4
1051	1539	1739	563	0.1111			363	0.0444			1339				851	0.4444	1.464	1.655	488	0.1111			688	0.3111	1.464	1.655	200	0.3778	1.464	1.655	3
1051	1539	1963	563	0.1905			139	0.1905			1115				627	0.2571	1.464	1.868	488	0.1238			912	0.3905	1.464	1.868	424	0.3905	1.464	1.868	3
1051	1739	1963	363	0.0333			139	0.0333			1515				827	0.1			688	0.2333	1.655	1.868	912	0.4333	1.655	1.868	224	0.5	1.655	1.868	3
1196	1358	1539	1034	0.381	1.135	1.287	853	0.1143			1177	-0.019			1015	0.4476	1.135	1.287	162	0.0476			343	0.381	1.135	1.287	181	0.181			3
1196	1358	1739	1034	0.6286	1.135	1.454	653	-0.0381			977	0.4952	1.135	1.454	815	0.5619	1.135	1.454	162	0.3619	1.135	1.454	543	0.2286			381	0.2952	1.135	1.454	5
1196	1358	1963	1034	0.7905	1.135	1.641	429	-0.0762			753	0.0571			591	0.1238			162	0.1905			767	0.2571	1.135	1.641	605	0.2571	1.135	1.641	3
1196	1358	2212	1034	0.7481	1.135	1.849	180	-0.0519			504	0.1481			342	0.0815			162	0.3481	1.135	1.849	1016	0.6815	1.135	1.849	854	0.1481			3
1196	1539	1739	853	0.4889	1.287	1.454	653	0.2889	1.287	1.454	1339				996	0.6889	1.287	1.454	343	0.0889			543	0.2222			200	0.4889	1.287	1.454	4
1196	1539	1963	853	0.0741			429	-0.3926			1115	0.1407			772	-0.0593			343	-0.1259			767	0.0741			424	0.0074			0
1196	1539	2212	853	0.5429	1.287	1.849	180	0.0762			866	0.1429			523	0.0095			343	0.4095	1.287	1.849	1016	0.2762	1.287	1.849	673	0.0762			3
1196	1739	1963	653	0.1556			429	0.0222			1515				972	0.2222			543	-0.0444			767	0.2889	1.454	1.641	224	0.3556	1.454	1.641	2
1196	1739	2212	653	0.0556			180	0.0556			1266				723	0.2556	1.454	1.849	543	0.0556			1016	0.3889	1.454	1.849	473	0.4556	1.454	1.849	3
1196	1963	2212	429	0.1222			180	0.0556			1714				947	0.2556	1.641	1.849	767	0.3222	1.641	1.849	1016	0.1889			249	0.7222	1.641	1.849	3
1358	1539	1739	1177	0.1333			977	0			1339	-0.0667			1158	0.3333	1.133	1.281	181	-0.0667			381	-0.2			200	0			1
1358	1539	1963	1177	0.7429	1.133	1.446	753	0.2095			1115	0.4095	1.133	1.446	934	0.5429	1.133	1.446	181	0.1429			605	0.3429	1.133	1.446	424	0.4095	1.133	1.446	5
1358	1539	2212	1177	0.7	1.133	1.629	504	0.0333			866	-0.0333			685	0.0333			181	0.3667	1.133	1.629	854	0.1667			673	0.0333			2
1358	1539	2489	1177	0.781	1.133	1.833	227	0.181			589	0.0476			408	0.2476	1.133	1.833	181	0.381	1.133	1.833	1131	0.581	1.133	1.833	950	0.381	1.133	1.833	5
1358	1739	1963	977	0.4667	1.281	1.446	753	0			1515				1134	0.4667	1.281	1.446	381	0.2			605	0.3333	1.281	1.446	224	0.5333	1.281	1.446	4
1358	1739	2212	977	0.4571	1.281	1.629	504	-0.0095			1266	0.4571	1.281	1.629	885	0.1238			381	0.2571	1.281	1.629	854	0.1905			473	0.2571	1.281	1.629	4
1358	1739	2489	977	0.3143	1.281	1.833	227	-0.019			989	0.181			608	0.1143			381	0.3143	1.281	1.833	1131	0.4476	1.281	1.833	750	0.181			3
1358	1963	2212	753	-0.0857			504	-0.1524			1714				1109	0.3143	1.446	1.629	605	-0.1524			854	-0.019			249	0.2476	1.446	1.629	2
1358	1963	2489	753	0.2889	1.446	1.833	227	0.0889			1437				832	0.2889	1.446	1.833	605	0.2222			1131	0.4222	1.446	1.833	526	0.3556	1.446	1.833	4
1358	2212	2489	504	0.1556			227	0.0222			1935				1081	0.0222			854	0.2222			1131	-0.0444			277	0.7556	1.629	1.833	1
1539	1739	1963	1339	0.4	1.13	1.276	1115	0.4	1.13	1.276	1515	0			1315	0.4	1.13	1.276	200	0.1333			424	0.2667	1.13	1.276	224	0.3333	1.13	1.276	5
1539	1739	2212	1339	0.6952	1.13	1.437	866	0.2286			1266	0.4286	1.13	1.437	1066	0.4286	1.13	1.437	200	0.4952	1.13	1.437	673	0.2952	1.13	1.437	473	0.4952	1.13	1.437	6
1539	1739	2489	1339	0.7333	1.13	1.617	589	0			989	0.2			789	0.3333	1.13	1.617	200	0.4	1.13	1.617	950	0.3333	1.13	1.617	750	0.2			4
1539	1739	2798	1339	0.8762	1.13	1.818	280	-0.0571			680	0.0095			480	-0.0571			200	0.4095	1.13	1.818	1259	0.6095	1.13	1.818	1059	0.3429	1.13	1.818	4
1539	1963	2212	1115	0.3667	1.276	1.437	866	0.0333			1714				1290	0.7667	1.276	1.437	424	0.2333	1.276	1.437	673	0.3667	1.276	1.437	249	0.4333	1.276	1.437	5
1539	1963	2489	1115	0.5333	1.276	1.617	589	-0.0667			1437	0.2667	1.276	1.617	1013	0.2			424	0.4667	1.276	1.617	950	0.1333			526	0.1333			3
1539	1963	2798	1115	0.6857	1.276	1.818	280	0.019			1128	0.219	1.276	1.818	704	0.0857			424	0.1524			1259	0.2857	1.276	1.818	835	0.219			3
1539	2212	2489	866	0.1556			589	0.2222			1935				1262	0.3556	1.437	1.617	673	0.4222	1.437	1.617	950	0.4222	1.437	1.617	277	0.8889	1.437	1.617	4
1539	2212	2798	866	0.2889	1.437	1.818	280	0.1556			1626				953	0.1556			673	0.2889	1.437	1.818	1259	0.2889	1.437	1.818	586	0.4889	1.437	1.818	4
1539	2489	2798	589	0.0222			280	0.0889			2180				1230	0.2889	1.617	1.818	950	0.0889			1259	-0.0444			309	0.3556	1.617	1.818	2
1739	1963	2212	1515	0.7048	1.129	1.272	1266	0.5048	1.129	1.272	1714	0.0381			1490	0.5048	1.129	1.272	224	0.3048	1.129	1.272	473	0.4381	1.129	1.272	249	0.6381	1.129	1.272	6
1739	1963	2489	1515	0.8167	1.129	1.431	989	0.2167			1437	0.4167	1.129	1.431	1213	0.2833	1.129	1.431	224	0.4167	1.129	1.431	750	0.4833	1.129	1.431	526	0.4167	1.129	1.431	6
1739	1963	2798	1515	0.8333	1.129	1.609	680	0.0333			1128	0.2333	1.129	1.609	904	0.2333	1.129	1.609	224	0.5	1.129	1.609	1059	0.5	1.129	1.609	835	0.1667			5
1739	2212	2489	1266	0.2111			989	0.2111			1935				1462	0.6778	1.272	1.431	473	0.1444			750	0.3444	1.272	1.431	277	0.6778	1.272	1.431	3
1739	2212	2798	1266	0.419	1.272	1.609	680	0.0857			1626	0.4857	1.272	1.609	1153	0.3524	1.272	1.609	473	0.419	1.272	1.609	1059	0.3524	1.272	1.609	586	0.4857	1.272	1.609	6
1739	2489	2798	989	0.0778			680	0.0778			2180				1430	0.3444	1.431	1.609	750	0.2111			1059	0.2111			309	0.6778	1.431	1.609	2
1963	2212	2489	1714	0.5714	1.127	1.268	1437	0.4381	1.127	1.268	1935	0.1714			1686	0.6381	1.127	1.268	249	0.4381	1.127	1.268	526	0.5048	1.127	1.268	277	0.7048	1.127	1.268	6
1963	2212	2798	1714	0.619	1.127	1.425	1128	0.019			1626	0.4857	1.127	1.425	1377	0.2857	1.127	1.425	249	0.219			835	0.1524			586	0.4857	1.127	1.425	4
1963	2489	2798	1437	0.5667	1.268	1.425	1128	0.1			2180				1654	0.5667	1.268	1.425	526	0.5	1.268	1.425	835	0.2333	1.268	1.425	309	0.5667	1.268	1.425	5
2212	2489	2798	1935	0.8	1.125	1.265	1626	0.5333	1.125	1.265	2180	0.4	1.125	1.265	1903	0.4667	1.125	1.265	277	0.5333	1.125	1.265	586	0.4667	1.125	1.265	309	0.4667	1.125	1.265	7
Min:					1.125	1.265			1.125	1.265			1.125	1.265			1.125	1.265			1.125	1.265			1.125	1.265			1.125	1.265	
Max:					1.527	1.946			1.474	1.868			1.299	1.868			1.671	1.917			1.671	1.946			1.671	1.888			1.713	1.946	

Table 3.6. Three-tone difference tones from Group A showing the discrimination indices for the seven difference tone classes investigated (low values white, high values dark red). The primary tone ratios are only displayed if the given difference tone is above threshold, and accordingly, the blank ratios represent difference tones below threshold. The rightmost column shows the total number of difference tones above threshold for the row's set of primary tones (maximum value is 7). The minimum and maximum f_2/f_1 and f_3/f_1 ratios in each class are in the final rows.

f1	f2	2f1-f2	D.I.%	f2/f1	f2-f1	D.I.%	f2/f1	Total
519	603	435	0.2667	1.162	84	0.6667	1.162	2
519	698	340	0.5333	1.345	179	0.7333	1.345	2
519	803	235	0.2	1.547	284	0.6667	1.547	2
519	921	117	0.2333	1.775	402	0.3667	1.775	2
603	698	508	0.1111		95	0.8444	1.158	1
603	803	403	0.4333	1.332	200	0.6333	1.332	2
603	921	285	0.5333	1.527	318	0.5333	1.527	2
603	1051	155	0.3	1.743	448	0.4333	1.743	2
698	803	593	0.2667	1.15	105	0.8	1.15	2
698	921	475	0.6	1.319	223	0.8	1.319	2
698	1051	345	0.1		353	0.3667	1.506	1
698	1196	200	0.6667	1.713	498	0.6	1.713	2
698	1358	38	-0.0444		660	-0.0444		0
803	921	685	0.6	1.147	118	0.8667	1.147	2
803	1051	555	0.7667	1.309	248	0.7667	1.309	2
803	1196	410	0.5333	1.489	393	0.4667	1.489	2
803	1358	248	0.2667	1.691	555	0.6	1.691	2
803	1539	67	0.4667	1.917	736	0.1333		1
921	1051	791	0.2667	1.141	130	0.8667	1.141	2
921	1196	646	0.7	1.299	275	0.9	1.299	2
921	1358	484	0.8	1.474	437	0.8	1.474	2
921	1539	303	0.0333		618	0.5	1.671	1
921	1739	103	0.2667	1.888	818	0		1
1051	1196	906	0.4333	1.138	145	0.8333	1.138	2
1051	1358	744	0.7	1.292	307	0.8333	1.292	2
1051	1539	563	0.7333	1.464	488	0.7333	1.464	2
1051	1739	363	0.5333	1.655	688	0.4	1.655	2
1051	1963	139	0.2667	1.868	912	0.2	1.868	2
1196	1358	1034	0.5667	1.135	162	0.7667	1.135	2
1196	1539	853	0.3667	1.287	343	0.7667	1.287	2
1196	1739	653	0.4667	1.454	543	0.7333	1.454	2
1196	1963	429	0.4	1.641	767	0.4667	1.641	2
1196	2212	180	0.3667	1.849	1016	0.2333	1.849	2
1358	1539	1177	0.7	1.133	181	0.8333	1.133	2
1358	1739	977	0.4667	1.281	381	0.8	1.281	2
1358	1963	753	0.4667	1.446	605	0.4667	1.446	2
1358	2212	504	0.0667		854	0.4	1.629	1
1358	2489	227	0.1778		1131	0.2444	1.833	1
1539	1739	1339	0.4667	1.13	200	0.7333	1.13	2
1539	1963	1115	0.6333	1.276	424	0.7667	1.276	2
1539	2212	866	0.3	1.437	673	0.4333	1.437	2
1539	2489	589	0.1		950	0.3667	1.617	1
1539	2798	280	0.1		1259	0.2333	1.818	1
1739	1963	1515	0.6333	1.129	224	0.7	1.129	2
1739	2212	1266	0.5778	1.272	473	0.5778	1.272	2
1739	2489	989	0.3333	1.431	750	0.6	1.431	2
1739	2798	680	0.0333		1059	0.3	1.609	1
1963	2212	1714	0.5333	1.127	249	0.8667	1.127	2
1963	2489	1437	0.5333	1.268	526	0.7333	1.268	2
1963	2798	1128	0.1333		835	0.4	1.425	1
2212	2489	1935	0.4667	1.125	277	0.4667	1.125	2
2212	2798	1626	0.5333	1.265	586	0.7333	1.265	2
2489	2798	2180	0.3	1.124	309	0.3667	1.124	2
Min:				1.124			1.124	
Max:				1.917			1.868	

Table 3.7. Two-tone difference tones from Group B showing the discrimination indices for the $2f_1$-f_2 CDT and f_2-f_1 QDT classes (low values white, high values dark red). The primary tone ratios are only displayed if the given difference tone is above threshold, and accordingly, the blank ratios represent difference tones below threshold. The rightmost column shows the total number of difference tones above threshold for the row's set of primary tones (maximum value is 2). The minimum and maximum ratios in each class are in the final rows.

f1	f2	f3	2f1-f2	D.I.%	f2/f1	f3/f1	2f1-f3	D.I.%	f2/f1	f3/f1	2f2-f3	D.I.%	f2/f1	f3/f1	f1+f2-f3	D.I.%	f2/f1	f3/f1	f2-f1	D.I.%	f2/f1	f3/f1	f3-f1	D.I.%	f2/f1	f3/f1	f3-f2	D.I.%	f2/f1	f3/f1	Total
519	603	698	435	0.15			340	0.15			508	0.0167			424	-0.1167			84	0.0833			179	0.6167	1.162	1.345	95	0.4833	1.162	1.345	2
519	603	803	435	-0.0286			235	0.1714			403	0.0381			319	0.3048	1.162	1.547	84	0.3714	1.162	1.547	284	0.2381	1.162	1.547	200	0.5048	1.162	1.547	4
519	603	921	435	0.3333	1.162	1.775	117	0.2	1.162	1.775	285	0.3333	1.162	1.775	201	0.4	1.162	1.775	84	0.6667	1.162	1.775	402	0.0667			318	0.2667	1.162	1.775	6
519	698	803	340	0.5111	1.345	1.547	235	0.1111			593				414	0.5111	1.345	1.547	179	0.1778			284	0.3778	1.345	1.547	105	0.5778	1.345	1.547	4
519	698	921	340	0.2095	1.345	1.775	117	0.0095			475	0.0095			296	0.5429	1.345	1.775	179	0.3429	1.345	1.775	402	0.0095			223	0.7429	1.345	1.775	4
519	803	921	235	0.1667			117	0.5	1.547	1.775	685				401	0.1667			284	0.3	1.547	1.775	402	0.0333			118	0.5	1.547	1.775	3
603	698	803	508	0.1917			403	0.0583			593	0.0583			498	0.1917			95	0.4583	1.158	1.332	200	0.4583	1.158	1.332	105	0.325	1.158	1.332	3
603	698	921	508	0.1048			285	0.1048			475	0.3048	1.158	1.527	380	0.4381	1.158	1.527	95	0.4381	1.158	1.527	318	0.4381	1.158	1.527	223	0.6381	1.158	1.527	5
603	698	1051	508	0.2381	1.158	1.743	155	0.1714			345	0.6381	1.158	1.743	250	0.1714			95	0.5714	1.158	1.743	448	0.1714			353	0.5048	1.158	1.743	4
603	803	921	403	0.2	1.332	1.527	285	0			685				485	0.4	1.332	1.527	200	0.4667	1.332	1.527	318	0.3333	1.332	1.527	118	0.7333	1.332	1.527	5
603	803	1051	403	0.181			155	0.2476	1.332	1.743	555	0.0476			355	0.7143	1.332	1.743	200	0.3143	1.332	1.743	448	0.1143			248	0.7143	1.332	1.743	4
603	921	1051	285	0.5667	1.527	1.743	155	-0.0333			791				473	0.2333	1.527	1.743	318	0.4333	1.527	1.743	448	0.1			130	0.7667	1.527	1.743	4
698	803	921	593	-0.0286			475	0.3048	1.15	1.319	685	0.0381			580	0.1714			105	0.0381			223	0.4381	1.15	1.319	118	0.3714	1.15	1.319	3
698	803	1051	593	0.1048			345	0.3714	1.15	1.506	555	0.0381			450	0.3714	1.15	1.506	105	0.3048	1.15	1.506	353	0.5048	1.15	1.506	248	0.4381	1.15	1.506	5
698	803	1196	593	0.0381			200	0.2381	1.15	1.713	410	0.4381	1.15	1.713	305	0.3048	1.15	1.713	105	0.5714	1.15	1.713	498	0.1714			393	0.5714	1.15	1.713	5
698	803	1358	593	0.1619			38	-0.0381			248	0.0286			143	0.4286	1.15	1.946	105	0.7619	1.15	1.946	660	0.0286			555	0.0286			2
698	921	1051	475	0.5111	1.319	1.506	345	0.1778			791				568	0.2444	1.319	1.506	223	0.3778	1.319	1.506	353	0.2444	1.319	1.506	130	0.7111	1.319	1.506	5
698	921	1196	475	0.5167	1.319	1.713	200	-0.0167			646	-0.0167			423	0.7167	1.319	1.713	223	0.5167	1.319	1.713	498	0.3167	1.319	1.713	275	0.7833	1.319	1.713	5
698	921	1358	475	0.3714	1.319	1.946	38	-0.0286			484	0.2381	1.319	1.946	261	0.1714			223	0.4381	1.319	1.946	660	-0.0286			437	0.4381	1.319	1.946	4
698	1051	1196	345	0.3778	1.506	1.713	200	-0.0889			906				553	0.4444	1.506	1.713	353	0.3111	1.506	1.713	498	0.4444	1.506	1.713	145	0.6444	1.506	1.713	5
698	1051	1358	345	0.2476	1.506	1.946	38	0.0476			744				391	0.7143	1.506	1.946	353	0.3143	1.506	1.946	660	0.1143			307	0.6476	1.506	1.946	4
698	1196	1358	200	0.2556	1.713	1.946	38	-0.0111			1034				536	0.5222	1.713	1.946	498	0.2556	1.713	1.946	660	-0.0111			162	0.6556	1.713	1.946	4
803	921	1051	685	0.1333			555	0.3333	1.147	1.309	791	0			673	0.0667			118	0.2	1.147	1.309	248	0.3333	1.147	1.309	130	0.4	1.147	1.309	4
803	921	1196	685	0.4095	1.147	1.489	410	0.2095	1.147	1.489	646	0.1429			528	0.6762	1.147	1.489	118	0.6762	1.147	1.489	393	0.7429	1.147	1.489	275	0.6762	1.147	1.489	6
803	921	1358	685	0.1524			248	0.0857			484	0.4857	1.147	1.691	366	0.4857	1.147	1.691	118	0.819	1.147	1.691	555	0.419	1.147	1.691	437	0.5524	1.147	1.691	5
803	921	1539	685	0.2083	1.147	1.917	67	-0.0583			303	0.1417			185	0.4083	1.147	1.917	118	0.875	1.147	1.917	736	0.1417			618	0.075			3
803	1051	1196	555	0.3778	1.309	1.489	410	0.2444	1.309	1.489	906				658	0.2444	1.309	1.489	248	0.6444	1.309	1.489	393	0.6444	1.309	1.489	145	0.8444	1.309	1.489	6
803	1051	1358	555	0.5524	1.309	1.691	248	0.619	1.309	1.691	744	0.0857			496	0.6857	1.309	1.691	248	0.419	1.309	1.691	555	0.2857	1.309	1.691	307	0.619	1.309	1.691	6
803	1051	1539	555	0.2333	1.309	1.917	67	0.1			563	0.3	1.309	1.917	315	0.3	1.309	1.917	248	0.3	1.309	1.917	736	-0.0333			488	0.3667	1.309	1.917	5
803	1196	1358	410	0.4095	1.489	1.691	248	0.0762			1034				641	0.4762	1.489	1.691	393	0.3429	1.489	1.691	555	0.4762	1.489	1.691	162	0.8095	1.489	1.691	5
803	1196	1539	410	0.5333	1.489	1.917	67	0.1333			853				460	0.7333	1.489	1.917	393	0.3333	1.489	1.917	736	-0.0667			343	0.6	1.489	1.917	4
803	1358	1539	248	0.1333			67	0.1333			1177				622	0.3333	1.691	1.917	555	0			736	0.0667			181	0.6	1.691	1.917	2
921	1051	1196	791	0.075			646	0.2083	1.141	1.299	906	0.1417			776	0.1417			130	0.1417			275	0.275	1.141	1.299	145	0.6083	1.141	1.299	3
921	1051	1358	791	0.0286			484	0.2952	1.141	1.474	744	0.0286			614	0.2952	1.141	1.474	130	0.5619	1.141	1.474	437	0.5619	1.141	1.474	307	0.6952	1.141	1.474	5
921	1051	1539	791	0.0667			303	0.2	1.141	1.671	563	0.0667			433	0.5333	1.141	1.671	130	0.8	1.141	1.671	618	0.5333	1.141	1.671	488	0.6	1.141	1.671	5
921	1051	1739	791	0.219	1.141	1.888	103	0.0857			363	0.3524	1.141	1.888	233	0.4857	1.141	1.888	130	0.8857	1.141	1.888	818	0.2857	1.141	1.888	688	0.2857	1.141	1.888	6
921	1196	1358	646	0.5889	1.299	1.474	484	0.3222	1.299	1.474	1034				759	0.1222			275	0.4556	1.299	1.474	437	0.4556	1.299	1.474	162	0.7889	1.299	1.474	5
921	1196	1539	646	0.6381	1.299	1.671	303	-0.0286			853	0.1048			578	0.5714	1.299	1.671	275	0.7048	1.299	1.671	618	0.2381	1.299	1.671	343	0.6381	1.299	1.671	5
921	1196	1739	646	0.4762	1.299	1.888	103	0.2762	1.299	1.888	653	0.2095	1.299	1.888	378	0.0762			275	0.6095	1.299	1.888	818	0.0762			543	0.4095	1.299	1.888	5
921	1358	1539	484	0.419	1.474	1.671	303	0.019			1177				740	0.219	1.474	1.671	437	0.419	1.474	1.671	618	0.4857	1.474	1.671	181	0.7524	1.474	1.671	5
921	1358	1739	484	0.4	1.474	1.888	103	0.1333			977				540	0.9333	1.474	1.888	437	0.4	1.474	1.888	818	0.1333			381	0.8667	1.474	1.888	4
921	1539	1739	303	0.0556			103	-0.0111			1339				721	0.2556	1.671	1.888	618	0.3222	1.671	1.888	818	-0.1444			200	0.7889	1.671	1.888	3
1051	1196	1358	906	0.019			744	0.019			1034	-0.0476			889	-0.1143			145	0.2857	1.138	1.292	307	0.2857	1.138	1.292	162	0.219	1.138	1.292	3
1051	1196	1539	906	0.1417			563	0.3417	1.138	1.464	853	0.075			708	0.5417	1.138	1.464	145	0.8083	1.138	1.464	488	0.675	1.138	1.464	343	0.7417	1.138	1.464	5
1051	1196	1739	906	0.2095	1.138	1.655	363	0.2762	1.138	1.655	653	0.2095	1.138	1.655	508	0.6095	1.138	1.655	145	0.8762	1.138	1.655	688	0.2762	1.138	1.655	543	0.2762	1.138	1.655	7
1051	1196	1963	906	0.2	1.138	1.868	139	0.5333	1.138	1.868	429	-0.0667			284	-0.2			145	0.4	1.138	1.868	912	0.2667	1.138	1.868	767	0.0667			4
1051	1358	1539	744	0.5667	1.292	1.464	563	0.0333			1177				870	0.4333	1.292	1.464	307	0.5667	1.292	1.464	488	0.3667	1.292	1.464	181	0.7	1.292	1.464	5
1051	1358	1739	744	0.4519	1.292	1.655	363	0.2519	1.292	1.655	977	0.0519			670	0.4519	1.292	1.655	307	0.7852	1.292	1.655	688	0.1852			381	0.7185	1.292	1.655	5
1051	1358	1963	744	0.419	1.292	1.868	139	0.219	1.292	1.868	753	0.019			446	0.219	1.292	1.868	307	0.5524	1.292	1.868	912	0.0857			605	0.419	1.292	1.868	5
1051	1539	1739	563	0.5111	1.464	1.655	363	0.0444			1339				851	0.2444	1.464	1.655	488	0.2444	1.464	1.655	688	0.3778	1.464	1.655	200	0.7778	1.464	1.655	5
1051	1539	1963	563	0.1619			139	-0.0381			1115				627	0.4286	1.464	1.868	488	0.0952			912	0.1619			424	0.7619	1.464	1.868	2
1051	1739	1963	363	0.2889	1.655	1.868	139	-0.0444			1515				827	0.2889	1.655	1.868	688	0.2222	1.655	1.868	912	0.0222			224	0.6889	1.655	1.868	4
1196	1358	1539	1034	0.0952			853	0.0952			1177	0.0286			1015	0.0952			162	0.4286	1.135	1.287	343	0.1619			181	0.5619	1.135	1.287	2
1196	1358	1739	1034	0.1143			653	0.3143	1.135	1.454	977	0.0476			815	0.381	1.135	1.454	162	0.6476	1.135	1.454	543	0.6476	1.135	1.454	381	0.781	1.135	1.454	5
1196	1358	1963	1034	0.1524			429	0.0857			753	0.219	1.135	1.641	591	0.1524			162	0.819	1.135	1.641	767	0.1524			605	0.2857	1.135	1.641	3
1196	1358	2212	1034	0.2815	1.135	1.849	180	0.1481			504	-0.1185			342	-0.0519			162	0.6815	1.135	1.849	1016	0.2815	1.135	1.849	854	0.2148	1.135	1.849	4
1196	1539	1739	853	0.4111	1.287	1.454	653	0.2778	1.287	1.454	1339				996	0.4111	1.287	1.454	343	0.7444	1.287	1.454	543	0.5444	1.287	1.454	200	0.8778	1.287	1.454	6
1196	1539	1963	853	0.1333			429	0			1115	0.0667			772	-0.1333			343	0			767	0.1333			424	0.1333			0
1196	1539	2212	853	0.619	1.287	1.849	180	0.1524			866	0.1524			523	0.4857	1.287	1.849	343	0.6857	1.287	1.849	1016	0.219	1.287	1.849	673	0.419	1.287	1.849	5
1196	1739	1963	653	0.1778			429	0.3111	1.454	1.641	1515				972	0.2444	1.454	1.641	543	0.1111			767	0.3111	1.454	1.641	224	0.8444	1.454	1.641	4
1196	1739	2212	653	0.2778	1.454	1.849	180	0.0111			1266				723	0.3444	1.454	1.849	543	0.6778	1.454	1.849	1016	0.1444			473	0.6778	1.454	1.849	4
1196	1963	2212	429	0.4111	1.641	1.849	180	0.0778			1714				947	0.2778	1.641	1.849	767	0.2111	1.641	1.849	1016	0.0111			249	0.8111	1.641	1.849	4
1358	1539	1739	1177	0.2476	1.133	1.281	977	0.2476	1.133	1.281	1339	-0.0857			1158	0.2476	1.133	1.281	181	0.3143	1.133	1.281	381	0.3143	1.133	1.281	200	0.4476	1.133	1.281	6
1358	1539	1963	1177	0.019			753	0.1524			1115	0.1524			934	0.419	1.133	1.446	181	0.8857	1.133	1.446	605	0.5524	1.133	1.446	424	0.4857	1.133	1.446	4
1358	1539	2212	1177	0.225	1.133	1.629	504	0.025			866	0.025			685	0.1583			181	0.425	1.133	1.629	854	0.0917			673	0.2917	1.133	1.629	3
1358	1539	2489	1177	0.381	1.133	1.833	227	-0.0857			589	0.1143			408	0.1143			181	0.781	1.133	1.833	1131	0.2476	1.133	1.833	950	0.181			3
1358	1739	1963	977	0.2889	1.281	1.446	753	0.0889			1515				1134	0.2222	1.281	1.446	381	0.0889			605	0.3556	1.281	1.446	224	0.6222	1.281	1.446	4
1358	1739	2212	977	0.6095	1.281	1.629	504	0.1429			1266	0.0762			885	0.3429	1.281	1.629	381	0.8095	1.281	1.629	854	0.2762	1.281	1.629	473	0.6095	1.281	1.629	5
1358	1739	2489	977	0.4476	1.281	1.833	227	-0.0857			989	0.181			608	0.1143			381	0.581	1.281	1.833	1131	0.1143			750	0.2476	1.281	1.833	3
1358	1963	2212	753	-0.1524			504	-0.219			1714				1109	-0.1524			605	-0.219			854	-0.019			249	0.581	1.446	1.629	1
1358	1963	2489	753	0.2111	1.446	1.833	227	0.0778			1437				832	0.6111	1.446	1.833	605	0.5444	1.446	1.833	1131	0.0111			526	0.6111	1.446	1.833	4
1358	2212	2489	504	0.2889	1.629	1.833	227	0.2222	1.629	1.833	1935				1081	0.0889			854	0.1556			1131	0.0222			277	0.7556	1.629	1.833	3
1539	1739	1963	1339	0.1905			1115	0.1238			1515	0.1238			1315	-0.0095			200	0.4571	1.13	1.276	424	0.3905	1.13	1.276	224	0.5905	1.13	1.276	3
1539	1739	2212	1339	0.4667	1.13	1.437	866	0.2	1.13	1.437	1266	0.1333			1066	0.1333			200	0.8	1.13	1.437	673	0.4667	1.13	1.437	473	0.6667	1.13	1.437	5
1539	1739	2489	1339	0.3083	1.13	1.617	589	0.0417			989	-0.025			789	0.1083			200	0.775	1.13	1.617	950	0.175			750	0.175			2
1539	1739	2798	1339	0.2667	1.13	1.818	280	0			680	0.0667			480	0.1333			200	0.8	1.13	1.818	1259	0.2667	1.13	1.818	1059	0.2	1.13	1.818	4
1539	1963	2212	1115	0.1778			866	0.0444			1714				1290	0.3111	1.276	1.437	424	0.3111	1.276	1.437	673	0.3111	1.276	1.437	249	0.7778	1.276	1.437	4
1539	1963	2489	1115	0.375	1.276	1.617	589	0.175			1437	0.0417			1013	0.2417	1.276	1.617	424	0.575	1.276	1.617	950	0.375	1.276	1.617	526	0.7083	1.276	1.617	5
1539	1963	2798	1115	0.3429	1.276	1.818	280	0.1429			1128	0.1429			704	0.0762			424	0.3429	1.276	1.818	1259	0.0762			835	0.5429	1.276	1.818	3
1539	2212	2489	866	0.1889			589	-0.0111			1935				1262	-0.0111			673	0.1889			950	0.0556			277	0.8556	1.437	1.617	1
1539	2212	2798	866	0.1778			280	0.1111			1626				953	0.3111	1.437	1.818	673	0.2444	1.437	1.818	1259	0.0444			586	0.5778	1.437	1.818	3
1539	2489	2798	589	0.1444			280	0.1444			2180				1230	0.1444			950	0.0778			1259	0.0778			309	0.4778	1.617	1.818	1
1739	1963	2212	1515	0.181			1266	0.181			1714	0.0476			1490	0.2476	1.129	1.272	224	0.3143	1.129	1.272	473	0.5143	1.129	1.272	249	0.4476	1.129	1.272	4
1739	1963	2489	1515	0.1333			989	0.2	1.129	1.431	1437	0.4	1.129	1.431	1213	0.3333	1.129	1.431	224	0.6667	1.129	1.431	750	0.5333	1.129	1.431	526	0.6	1.129	1.431	6
1739	1963	2798	1515	0.275	1.129	1.609	680	0.075			1128	0.075			904	0.1417			224	0.6083	1.129	1.609	1059	0.1417			835	0.275	1.129	1.609	3
1739	2212	2489	1266	0.2556	1.272	1.431	989	-0.0778			1935				1462	0.3889	1.272	1.431	473	0.4556	1.272	1.431	750	0.3889	1.272	1.431	277	0.8556	1.272	1.431	5
1739	2212	2798	1266	0.4571	1.272	1.609	680	0.0571			1626	0.0571			1153	0.1905			473	0.6571	1.272	1.609	1059	0.1905			586	0.5238	1.272	1.609	3
1739	2489	2798	989	-0.0111			680	-0.0778			2180				1430	0.1889			750	0.3222	1.431	1.609	1059	0.3222	1.431	1.609	309	0.7222	1.431	1.609	3
1963	2212	2489	1714	0.1429			1437	0.3429	1.127	1.268	1935	0.1429			1686	0.2095	1.127	1.268	249	0.4762	1.127	1.268	526	0.5429	1.127	1.268	277	0.3429	1.127	1.268	5
1963	2212	2798	1714	-0.0095			1128	-0.0095			1626	0.2571	1.127	1.425	1377	0.1238			249	0.7238	1.127	1.425	835	0.3238	1.127	1.425	586	0.7238	1.127	1.425	4
1963	2489	2798	1437	0.2889	1.268	1.425	1128	0.0222			2180				1654	0.2889	1.268	1.425	526	0.2222	1.268	1.425	835	0.3556	1.268	1.425	309	0.6889	1.268	1.425	5
2212	2489	2798	1935	0.0074			1626	0.1407			2180	-0.0593			1903	0.2074	1.125	1.265	277	0.5407	1.125	1.265	586	0.7407	1.125	1.265	309	0.6074	1.125	1.265	4
Min:					1.129	1.281			1.127	1.268			1.127	1.425			1.125	1.265			1.125	1.265			1.125	1.265			1.125	1.265	
Max:					1.713	1.946			1.629	1.888			1.319	1.946			1.713	1.946			1.713	1.946			1.506	1.888			1.713	1.946	

Table 3.8. Three-tone difference tones from Group B showing the discrimination indices for the seven difference tone classes investigated (low values white, high values dark red). The primary tone ratios are only displayed if the given difference tone is above threshold, and accordingly, the blank ratios represent difference tones below threshold. The rightmost column shows the total number of difference tones above threshold for the row's set of primary tones (maximum value is 7). The minimum and maximum f_2/f_1 and f_3/f_1 ratios in each class are in the final rows.

Two-Tone Difference Tones

All of the two-tone distortion products investigated in this study are reported in Figure 3.1, with each cluster representing the $2f_1$-f_2 and the f_2-f_1 distortion product f_1 and f_2/f_1 ratio for both Group A and Group B. The dots on the bottom of each cluster are Group A, the top row of each cluster is Group B, the left column of each cluster is the $2f_1$-f_2 distortion product, and the right column is the f_2-f_1 distortion product. The blue dots represent the f_1 and f_2/f_1 ratio for distortion products above threshold, and the red dots are below threshold. The upper limit is mostly visible, although the lower ratio threshold is not.

$2f_1$-f_2

In Group A, eight f_1 primary tone groups show the maximum ratio limit within the octave, however the minimum ratio is not detectable in any of the primary tone combinations. There are two cases of irregularities with non-detection in the middle of the ratio range. While the ratio changes across the f_1 frequency range, the minimum f_2/f_1 ratio is 1.124 for any primary tone combination and the maximum ratio for any combination is 1.775.

The results in Group B show the maximum ratio for five f_1 primary tone groups, the minimum ratio is detected in one f_1 group, and there are two irregularities. In six of the f_1 groups, every distortion product tested is above threshold, hence the boundaries are not found. The overall minimum f_2/f_1 ratio detected is 1.124, and the maximum ratio is 1.917.

f_2-f_1

In the second class of the two-tone primaries tested, the results in Group A show four cases of the maximum ratio detected, two irregularities of non-detection, and no cases of the minimum ratio. The smallest f_2/f_1 ratio is 1.124 and the largest is nearly an octave at 1.946, which occurs for the f_1 primary tone 698 Hz.

Nearly all distortion products are above threshold in Group B. There are ten cases with no lower or upper ratio limit, and only three cases with an upper limit detected. The three distortion products that comprise the upper limit are also the only cases of distortion products below threshold in this class and group. The overall minimum ratio is 1.124, and the overall maximum ratio detected is 1.868.

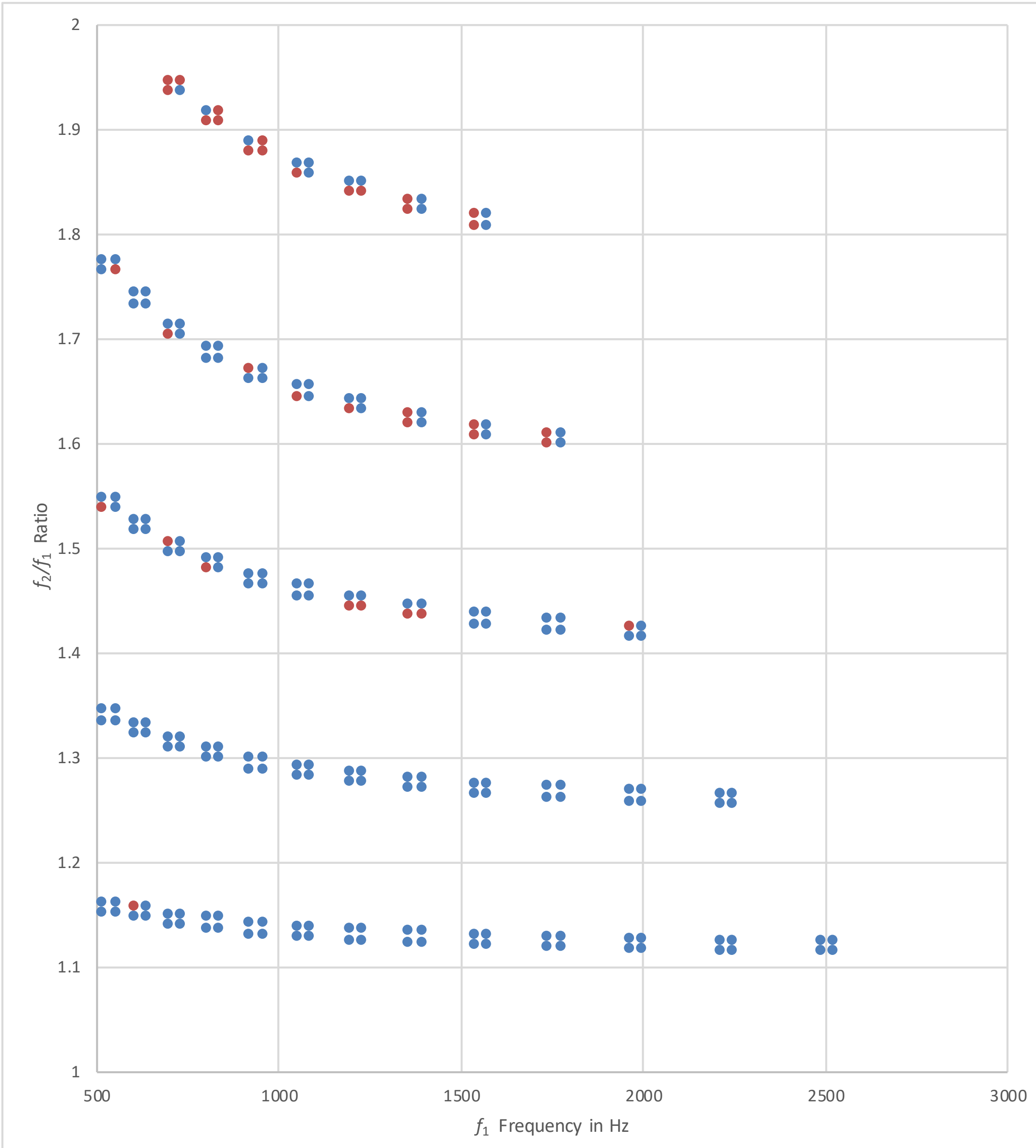


Figure 3.1. Two-tone results, with each square cluster of dots representing a single entry. In each cluster, the bottom two dots are Group A, the top row Group B, the left two dots are the $2f_1$-f_2 CDTs and the two dots on the right are the f_2-f_1 QDTs. The blue dots are difference tones above threshold and the red dots are below threshold. The f_1 frequency and f_2/f_1 ratios are slightly offset for display purposes.

Three-Tone Difference Tones

With the addition of a third primary tone comes nested subgroup ratio ranges within each f_1 difference tone class. For example, the f_1 primary tone of 519 is divided into three subgroups delineated by the thin borders in table 3.6 and table 3.8: those containing a f_2 frequency of 603 Hz and an increasing f_3 primary, another group with a f_2 of 698 Hz and an increasing f_3 range, and a single maximum primary tone combination of f_2=803 Hz, f_3=921 Hz. By viewing the results through subgroups, one can see that instead of irregularities of detection and non-detection, there are true boundary limits for the primary tone ratios as the respective f_2 and f_3 tones increase. One must be careful when interpreting the irregular sequences of non-detection as to not confuse the non-detection with the maximum range for a subgroup of primaries in that particular f_1 class. To illustrate this, let us consider an example of the Group B $2f_1$-f_2 difference tone in the subgroup of f_1=803 Hz and f_2=921 Hz. We first see the lowest ratio boundary through the non-detection of the 803 Hz, 921 Hz, 1051 Hz case. As f_3 increases, we find detection at 803 Hz, 921 Hz, 1196 Hz, making this condition the lowest detectable distortion product in the subgroup. As the f_3 increases, we once again find non-detection at the next ratio step of 803 Hz, 921 Hz, 1358 Hz. This set of primaries could be misconstrued as the upper limit in the subgroup, however we once again find detection at the final step in the subgroup 803 Hz, 921 Hz, 1539 Hz. The nested non-detection at 803 Hz, 921 Hz, 1358 Hz is therefore not an upper boundary but rather an example of the irregular behavior of difference tone detection.

The three-tone distortion products are represented in the three-dimensional Figure 3.2. The dots represent difference tones as a function of their f_1 primary tone frequencies and f_2/f_1 and f_3/f_1 primary tone ratios. The blue dots represent resulting difference tones above threshold, and the red dots represent difference tones below threshold. The Group A results are in the foreground, and the Group B results are offset to the row behind/above. The three-tone difference tone classes are listed in the following order from left to right: $2f_1$-f_2, $2f_1$-f_3, $2f_2$-f_3, f_1+f_2-f_3, f_2-f_1, f_3-f_1, f_3-f_2, which is the same order as in table 3.6 and table 3.8. The f_1 frequency is on the x-axis, with each seven-tone cluster representing the same f_1 for all seven three-tone difference tone classes. The y-axis is the primary tone ratio between f_2/f_1 and the z-axis is the primary tone ratio between f_3/f_1. The cases with gaps

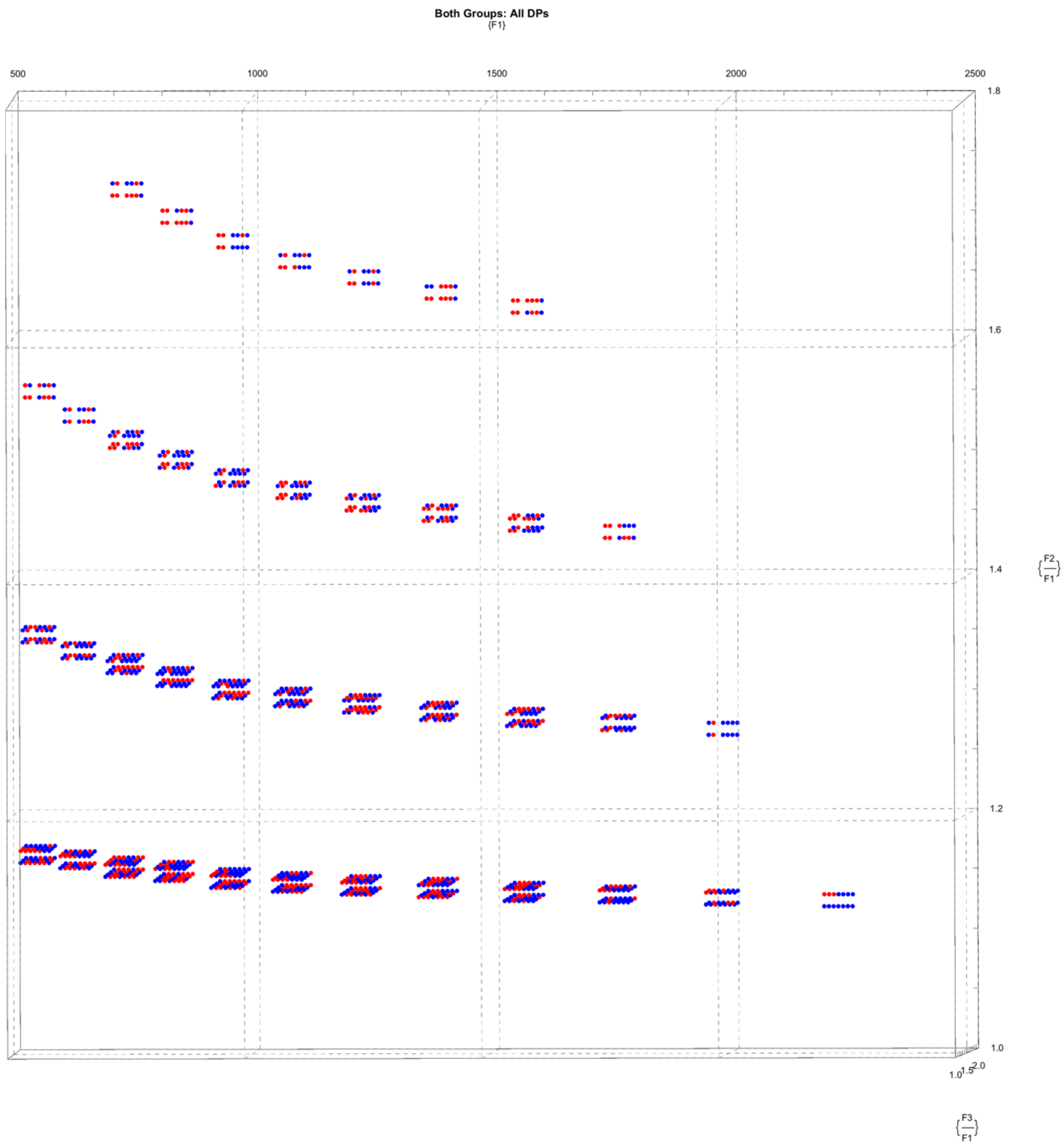


Figure 3.2. *Three-tone difference tones, offset to include both groups and all seven classes. The blue dots are difference tones above threshold, and red dots below threshold. In each cluster, the order of the difference tones is from left to right: $2f_1$-f_2, $2f_1$-f_3, $2f_2$-f_3, f_1+f_2-f_3, f_2-f_1, f_3-f_1, f_3-f_2. The row of dots on the bottom/front of each cluster are Group A and the top/back row are Group B. The x-axis is the f_1 frequency, the y-axis is the f_2/f_1 ratio, and the z-axis is the f_3/f_1 ratio.*

(missing dots) are instances of primary tone relationships that did not meet inclusion criteria and therefore were not tested. The three or four tiered groupings are the subgroups within each f_1 group. For example, the primary tones with f_1= 519 Hz have three subgroups: on the bottom is the subgroup with an f_2 of 603 Hz, the middle tier is the subgroup with an f_2 of 698 Hz, and the top tier is the subgroup with a f_2 of 803 Hz. Within each of these subgroups, the rising f_3/f_1 relationship is shown by the z-axis height.

$2f_1$-f_2

As illustrated in table 3.6, Group A shows strong detection. Maximum ratio boundaries are clear in nine main blocks of primaries with the same f_1 frequency. In the subgroups, there are five cases of the minimum ratio boundary, two cases of the maximum ratio boundary (not mutually exclusive with the f_1 block maximum), and one case of an irregular non-detection nested between cases of detection. No subgroups show both minimum and maximum ratio limits. The minimum f_2/f_1 ratio is 1.125 and minimum f_3/f_1 ratio is 1.265 for all Group A three-tone difference tones investigated. The overall highest f_2/f_1 ratio for this specific distortion product is 1.527, and the overall highest f_3/f_1 ratio in this condition is 1.946.

Table 3.8 shows the results for Group B. There are six cases of clear maximum boundaries in blocks with the same f_1 frequency. In the subgroups, there are 11 cases of the minimum ratio boundaries, two cases of the maximum boundaries, and three cases of irregular non-detection. Like Group A, there are no cases in the subgroups with clear upper and lower boundaries. The overall minimum f_2/f_1 ratio is 1.129 and the minimum f_3/f_1 ratio is 1.281. The overall maximum f_2/f_1 ratio is 1.1713 and the maximum f_3/f_1 ratio is 1.946.

$2f_1$-f_3

In Group A, detection is sparse for the $2f_1$-f_3 difference tone, with as few as zero detection in two of the primary tone combination blocks with the same f_1 frequency (f_1=603 Hz and f_1=1358 Hz). This difference tone class is one of the three with a mean discrimination index below threshold. There are nine cases of clear maximum ratio limits in blocks with the same f_1 frequency. In the subgroups, there are two cases with a minimum ratio limit, nine cases with a maximum limit, and no cases of irregular non-detection. There are also no subgroups with both minimum and maximum ratio limits. The overall minimum ratios are as previously stated in the $2f_1$-f_2 class results, the overall maximum f_2/f_1 ratio is 1.474, and the overall maximum f_3/f_1 ratio is 1.868.

The mean discrimination index for Group B is also below threshold, although there is more detection of difference tones in this group. There are ten cases of clear maximum ratio boundaries in blocks with the same f_1 frequency. Seven subgroups show the minimum ratio limit, 11 show the maximum ratio boundary, and there is one case of irregular non-detection in the middle of the subgroup. The collected data shows the full upper and lower ratio limit in three subgroup cases (not mutually exclusive with the other boundaries reported in this group). The overall minimum f_2/f_1 ratio is 1.127 and the minimum f_3/f_1 ratio is 1.268. The overall maximum f_2/f_1 ratio is 1.1629 and the maximum f_3/f_1 ratio is 1.888.

$2f_2$-f_3

The $2f_2$-f_3 difference tone class is the second of three Group A distortion products with a mean discrimination index below threshold. As a result, detection is sporadic and there is one f_1 primary tone group that shows no detection (f_1=803 Hz). The maximum ratio boundaries are clear in six of the blocks with the same f_1 frequency. The subgroups show the minimum threshold in 11 cases, the maximum threshold in eight cases, and there are no cases of irregular detection. Six of the subgroups contain both the minimum and maximum ratio boundaries. The overall maximum f_2/f_1 ratio is 1.299, and the overall maximum f_3/f_1 ratio is 1.868.

Like its counterpart group, the mean discrimination index in Group B is also below threshold and has few cases of detection. There are three blocks of primaries with the same f_1 frequency that show no detection (f_1=1358 Hz, f_1=1539 Hz, and the single case block f_1=2212 Hz). The maximum ratio boundaries are clear in five blocks with the same f_1. In the subgroups, there are 12 cases of the minimum threshold detected, five cases of the maximum threshold, and no cases of irregular non-detection. Five subgroups show both the minimum and maximum ratio boundaries. The overall minimum f_2/f_1 ratio is 1.127 and the minimum f_3/f_1 ratio is 1.425. The overall maximum f_2/f_1 ratio is 1.319 and the maximum f_3/f_1 ratio is 1.946.

$f_1+f_2-f_3$

In Group A, there are four cases of the maximum ratio boundaries for entire blocks with the same f_1 frequency. The subgroups show three cases of the minimum boundary, 16 cases of the maximum boundary, and one case of irregular non-detection. Two subgroups show both the upper and lower boundaries. The overall maximum f_2/f_1 ratio is 1.671, and the overall maximum f_3/f_1 ratio is 1.917.

There are also four maximum boundaries for entire f_1 blocks in Group B. Ten subgroups have a minimum threshold, 11 subgroups have a maximum threshold, and one subgroup contains irregular non-detection. Four subgroups have both upper and lower boundaries. The overall minimum f_2/f_1 ratio is 1.125 and the minimum f_3/f_1 ratio is 1.265. The overall maximum f_2/f_1 ratio is 1.713 and the maximum f_3/f_1 ratio is 1.946.

f_2-f_1

The f_2-f_1 difference tone is the third of the three distortion products investigated in Group A with a mean discrimination index below threshold. The three-tone f_2-f_1 difference tone shows detection is equal to or greater than twice the detection in the other two below-threshold classes in Group A. Seven blocks with the same f_1 primary show maximum ratio boundaries. In the subgroups, there are 11 cases of the minimum ratio boundary, eight cases with the maximum boundary, and one subgroup with irregular non-detection. Three of the subgroups show both the minimum and maximum ratio boundaries. The overall maximum f_2/f_1 ratio is 1.671, and the overall maximum f_3/f_1 ratio is 1.946.

The mean discrimination index in Group B is above threshold and the group contains many more detections than its counterpart. There are three cases of clear maximum ratio boundaries in primary tone blocks with the same f_1 frequency. Eight subgroups show the minimum ratio boundaries, one subgroup contains the maximum boundary, and there is one case with irregular non-detection. No subgroups contain both the upper and lower boundaries. It is interesting to note the low number of cases with maximum ratio boundaries, and the large number of detections in general. The overall minimum f_2/f_1 ratio is 1.125 and the minimum f_3/f_1 ratio is 1.265. The overall maximum f_2/f_1 ratio is 1.713 and the maximum f_3/f_1 ratio is 1.946.

f_3-f_1

Eight primary tone blocks with the same f_1 frequency show the maximum ratio boundary in Group A. In the subgroups, there are eight cases showing the minimum threshold, 12 cases showing the maximum threshold, and four cases with irregular non-detection. Four subgroups contain both the upper and lower ratio boundaries. The overall maximum f_2/f_1 ratio is 1.671, and the overall maximum f_3/f_1 ratio is 1.888.

In Group B, the results show nine cases of maximum ratio boundaries within a block of primary tones with the same f_1. In the subgroups, there is one case of the minimum ratio boundary, 19 cases of the maximum ratio boundary, and four cases of irregular non-detection. No subgroups contain both maximum and minimum ratios. The overall minimum f_2/f_1 ratio is 1.125 and the minimum f_3/f_1 ratio is 1.265. The overall maximum f_2/f_1 ratio is 1.506 and the maximum f_3/f_1 ratio is 1.888.

f_3-f_2

In the final distortion product examined, Group A does not exhibit any cases of the maximum boundary within a single block of primary tones with the same f_1 frequency. In the subgroups, there are seven cases of the minimum ratio boundaries, 13 cases of the maximum ratio boundaries, four cases of irregular non-detection, and four cases with both minimum and maximum ratio boundaries. The overall maximum f_2/f_1 ratio is 1.713, and the overall maximum f_3/f_1 ratio is 1.946.

Likewise, Group B also does not contain any cases of the maximum ratio boundaries in a block of primaries with the same f_1 frequency, making the f_3-f_2 the only difference tone class in the study with this characteristic. Otherwise, detection is high in this group, which is another similarity with Group A. There are no subgroups with minimum ratio boundaries, four subgroups with a maximum boundary, two subgroups with irregular non-detection, and no subgroups with both minimum and maximum ratio boundaries. The overall minimum f_2/f_1 ratio is 1.125 and the minimum f_3/f_1 ratio is 1.265. The overall maximum f_2/f_1 ratio is 1.713 and the maximum f_3/f_1 ratio is 1.946.

3.4 DISCUSSION AND ANALYSIS

The results indicate all of the difference tone classes contain primary tone configurations that evoke perceivable and statistically reliable distortion products across the frequency spectrum tested. Each difference tone class shows unique upper and/or lower primary tone ratio boundaries. From a single trial of primary tones, participants can perceive as many as all seven three-tone difference tone classes and both of the two-tone difference tones. Although small in magnitude, there are statistically significant differences between Group A and Group B. As reported in table 3.4, the mean discrimination index for all difference tones is 0.2913 in Group A and 0.3344 in Group B. The difference between these means is 0.0431, which is significant as demonstrated by a two-tail t-test (t = -4.372, p < 1.41e-5). We will first examine these group differences before unpacking the rest of the collected data.

Group A has 715 trials with a probe tone at the level of L_p=L_1-20 dB, and an additional subgroup of 36 duplicate trials with a probe tone at L_p=L_1-30 dB. Group B has 715 trials with a probe tone of L_p=L_1-30 dB, and a subgroup of 36 of the same duplicate trials with a probe tone of L_p=L_1-20 dB. The 36 trials in the subgroups are compared against the same 36 trials in the main groups. Using this data, there are four variations to compare: Group B main stimuli vs Group A subgroup (L_p=L_1-30 dB vs L_p=L_1-30 dB), Group B main stimuli vs Group B subgroup (L_p=L_1-30 dB vs L_p=L_1-20 dB), Group A main stimuli vs Group B subgroup (L_p=L_1-20 dB vs L_p=L_1-20 dB), and Group A main stimuli vs Group A subgroup (L_p=L_1-20 dB vs L_p=L_1-30 dB).

First let us compare the two sets of trials with a L_p=L_1-30 dB probe tone, which is the condition of the Group B main stimuli and Group A subgroup. The mean discrimination index for the Group B main trials is 0.474, and 0.537 for the Group A subgroup. The difference of the means (0.063) is statistically significant as demonstrated by a two-tail t-test (t = -3.705, p < 0.0007). This comparison shows Group A is better at detection than Group B when the probe is L_p=L_1-30 dB.

For the second comparison we will examine the entirely within-Group B condition, which is comparing the main Group B stimuli that has a probe at L_p=L_1-30 dB against the Group B subgroup that has the alternate probe level of L_p=L_1-20 dB. The mean discrimination index for the main stimuli with the L_p=L_1-30 dB probe is 0.474 and the mean discrimination index for the subgroup stimuli with the L_p=L_1-20 dB probe is 0.446. The difference of the means (0.028) is not statistically significant as demonstrated by a two-tail t-test ($t = 1.054$, $p > 0.2989$). This comparison indicates there is not a significant difference within Group B when the probe tone is L_p=L_1-30 dB or L_p=L_1-20 dB.

For the third comparison, let us examine the two cases with a L_p=L_1-20 dB probe tone, which is specifically the main stimuli from Group A and the Group B subgroup. The mean discrimination index for the Group A stimuli is 0.467 and the mean discrimination index for the Group B subgroup is 0.446. The difference between the means (0.021) is not statistically significant as demonstrated by a two-tail t-test ($t = 0.725$, $p > 0.4731$). These results suggest Group A and Group B do not differ when the probe tone is L_p=L_1-20 dB.

For the fourth comparison, we examine the entirely within-Group A condition, which is the main stimuli in Group A with a probe tone at L_p=L_1-20 dB and the Group A subgroup with the alternate probe level of L_p=L_1-30 dB. The mean discrimination index of the main stimuli with a probe at L_p=L_1-20 dB is 0.467 and the mean discrimination index of the subgroup with a probe at L_p=L_1-30 dB is 0.537. The difference between the means (0.070) is not statistically significant as demonstrated by a two-tail t-test (t =-1.898, $p >$ 0.0660), however it is approaching statistical significance (and the results are significant with a one tail t-test). While there is not a significant difference within Group A when the probe tone is L_p=L_1-20 dB or L_p=L_1-30 dB, the results suggest a trend that Group A detects difference tones better with a probe tone of L_p=L_1-30 dB.

Overall, these results show that Group A performs the detection task better than Group B when the probe tone is L_p=L_1-30 dB, but there is not a statistically significant difference between groups when the probe tone is L_p=L_1-20 dB. Examining the within-group data, the results show that Group B does not perform the task better with any specific probe tone level. Group A exhibits a trend for better detection with a probe tone at L_p=L_1-30 dB, however the effect is not large and not statistically significant. It appears the L_p=L_1-30 dB probe level is slightly better for the detection of beats, perhaps because the louder L_p=L_1-20 dB probe may mask some difference tones. This may also be the reason why Group A is better with the L_p= L_1-30 dB probe tone than Group B: if the main probe tone amplitude level for Group A is too loud and makes detection more difficult, Group A excels at the task in the relatively few cases when the probe tone is at a lower amplitude. The probe tone in Group B is usually L_p=L_1-30 dB, so the subjects do not show a large change when they receive a louder probe tone at L_p=L_1-20 dB in the 36 subgroup cases as their mean detection only decreases by 0.028.

Rank	Group A	Group B
1	**f_2-f_1 (2t) (0.4548)**	**f_2-f_1 (2t) (0.5715)**
2	2f_1-f_2 (2t) (0.4397)	f_3-f_2 (0.5564)
3	2f_1-f_2 (0.4127)	f_2-f_1 (0.4586)
4	f_3-f_2 (0.3504)	2f_1-f_2 (2t) (0.4017)
5	**f_1+f_2-f_3 (0.3141)**	**f_1+f_2-f_3 (0.2949)**
6	**f_3-f_1 (0.2438)**	**f_3-f_1 (0.2680)**
7	[f_2-f_1] (0.2199)	2f_1-f_2 (0.2615)
8	[2f_2-f_3] (0.1726)	[2f_1-f_3] (0.1253)
9	[2f_1-f_3] (0.1011)	[2f_2-f_3] (0.1188)

Table 3.9. Ranking of the nine difference tones by mean discrimination index (in parenthesis). The two-tone primaries are labeled as "2t," and the difference tone classes below threshold are in square brackets. The entries in bold indicate the difference tone class is at the same rank in both groups.

Moving on from the group differences, next we will analyze the nine classes of difference tones by ranking their mean discrimination indices. As overviewed in the Results, six classes of difference tones tested in Group A and seven classes in Group B have mean discrimination indices above threshold. The magnitude of the mean discrimination index shows the strength of detection of that particular distortion product class. Table 3.9 shows each distortion product class by rank of mean discrimination index. The two-tone variety are labeled as "(2t)." The difference tone classes in square brackets have a mean discrimination index below threshold. In both groups the class with the highest mean discrimination index is the two-tone f_2-f_1 difference tone. It is interesting to note that along with the two-tone f_2-f_1 difference tone, the f_1+f_2-f_3, and the f_3-f_1 distortion products also share the same rank in both groups. For clarity, difference tone classes that share the same rank are represented in bold text in the table. The mean discrimination index for the three-tone f_2-f_1 difference tone is close to but below threshold in Group A, and the third highest in Group B. If the data between these groups were combined, the three-tone f_2-f_1 difference tone would be above threshold. The Pearson correlation shows the degree of linear relationship between two rankings. The value varies from -1 to +1, with 1 showing a completely positive linear relationship, the opposite negative linear correlation with -1, and 0 showing no linear correlation. When comparing the groups, the order of mean discrimination indices shows some correlation (Pearson r_{xy} = 0.6631787).

Looking within each distortion product class, an embedded ranking emerges when examining the number of individual difference tones above threshold. Table 3.10 shows the ranking of each difference tone class by the percentage of trials above threshold. Again, the distortion products with a mean discrimination index below threshold are in square brackets, and the distortion products that are equal in position between groups are in bold text. All distortion products are the three-tone variety unless specified as "(2t)." In Group A, the two-tone distortion product f_2-f_1 was the most detectable with 89% of the individual distortion products above threshold. The remaining Group A difference tone detection percentages are listed from most to least: the three-tone 2f_1-f_2 at 75%, the f_3-f_2 at 71%, then the two-tone 2f_1-f_2 with 68%, followed by the f_1+f_2-f_3 difference tone at 67%, the f_3-f_1 with 52%, and the three classes with mean discrimination indices below threshold complete the ranking with f_2-f_1 at 43%, 2f_2-f_3 with 35%, and finally the 2f_1-f_3 with 16% detection above threshold. The distortion product with the most detection above threshold in Group B is the two-tone f_2-f_1 at 94% detection and it is nearly tied with f_3-f_2 at 93% detection. The remaining difference tones in Group B are ranked as follows: the three-tone f_2-f_1 difference tone at 86%, the two-tone 2f_1-f_2 difference tone at 81%, the f_1+f_2-f_3 at 65%, and nearly tied is the three-tone 2f_1-f_2 at 59%, the f_3-f_1 difference tone at 58%, and the remaining two with mean discrimination indices below threshold are 2f_1-f_3 at 30%, and 2f_2-f_3 with 23% detection.

When examining the percentage of detection within each difference tone class, it is interesting to note the relationship between the groups. The two-tone f_2-f_1 is the difference tone with the highest percent of detection in both groups, and the two-tone 2f_1-f_2 as well as the f_1+f_2-f_3 difference tone classes share the same rank position in their group order (fourth and fifth respectively). It is also notable the f_3-f_1 difference tone is the last type in both groups in the category of difference tones with a mean discrimination index above threshold. There is correlation between the groups for the percent of detection (Pearson r_{xy}

= 0.7369228). Comparing the same class of difference tones between groups, the percentages in Group B are better than Group A in all but three cases (both the three and two-tone $2f_1$-f_2, and $2f_2$-f_3).

Rank	Group A	Group B
1	**f_2-f_1 (2t) (89%)**	**f_2-f_1 (2t) (94%)**
2	$2f_1$-f_2 (75%)	f_3-f_2 (93%)
3	f_3-f_2 (71%)	f_2-f_1 (86%)
4	**$2f_1$-f_2 (2t) (68%)**	**$2f_1$-f_2 (2t) (81%)**
5	**f_1+f_2-f_3 (67%)**	**f_1+f_2-f_3 (65%)**
6	f_3-f_1 (52%)	$2f_1$-f_2 (59%)
7	[f_2-f_1] (43%)	f_3-f_1 (58%)
8	[$2f_2$-f_3] (35%)	[$2f_1$-f_3] (30%)
9	[$2f_1$-f_3] (16%)	[$2f_2$-f_3] (23%)

Table 3.10. Ranking of the nine difference tones by the percent of trials above threshold (in parenthesis). The two-tone difference tones are labeled "2t," and the entries in square brackets signify difference tone classes that have mean discrimination indices below threshold. The entries in bold indicate the class is at the same rank in both groups.

Within each group, there is high correlation between the ranking order of mean discrimination indices and the ranking order of difference tone class percent of detection (Group A Pearson r_{xy} = 0.9565163, and Group B Pearson r_{xy} = 0.9731883). There is also high correlation across group categories, such as comparing the mean discrimination index order for Group A to the percent of detection in Group B (Pearson r_{xy} = 0.7137043). The opposite case also has high correlation when comparing the Group B mean discrimination index ranking to the Group A ranking of percent of detection (Pearson r_{xy} = 0.7049355).

Comparing evaluations thus far, Group B has more difference tone classes with mean discrimination indices above threshold, has higher mean discrimination indices in general, and has higher percentages of detection than Group A. Based on this data, it appears the probe tone at a level of L_p=L_1-30 dB yields stronger results. However, this observation is complicated by Group A's superior performance with a L_p=L_1-30 dB probe tone compared to Group B's performance with the same probe level. This may be due to the previously suggested reason that the L_p=L_1-30 dB probe tones stand out better in Group A for they are in a minority of trials. Overall, it seems safe to suggest the L_p=L_1-30 dB probe tone produces more salient beats with the distortion products. The louder probe tone level of L_p=L_1-20 dB allows for the detection of beats between the probe and the distortion product, but a L_p=L_1-30 dB probe produces even more beats. That said, certain difference tone classes have higher discrimination indices with a L_p=L_1-20 dB probe tone, and a subset of those also have a higher percentage of detection at the same level. While the L_p=L_1-30 dB probe tone may seem advantageous, it is worth considering both groups as separate entities for these reasons.

Looking deeper within each of the difference tone classes, we can find the approximate primary tone ratio boundaries for above-threshold difference tones across the tested frequency spectrum. Specifically, we determine the range from the minimum and maximum f_2/f_1 and f_3/f_1 primary tone ratios that elicit above-threshold difference tones. This data is produced in all nine classes, regardless of the state of the mean discrimination index, however we are not always able to determine boundary limits.

As seen in table 3.5 and table 3.7, the blanks (indicating below-threshold difference tones) reveal the upper f_2/f_1 primary tone boundary for many cases of the two-tone $2f_1$-f_2 and two-tone f_2-f_1 difference tones. The lower primary tone ratio boundary is generally not found. Given the primary tones are at ERB center frequencies, the lowest possible f_2/f_1 ratio is 1.124, therefore the ERB scale rendering for the primary tones does not appear to have the resolution needed to find the lowest ratio threshold for the two-tone difference tones.

The three-tone conditions are expectedly more complex, however many of the primary tone ratio boundaries emerge when evaluating the nested subgroups as described in the Results section. Table 3.6 and table 3.8 show the minimum and maximum primary tone ratios that yield difference tones above threshold. The data generally shows approximate lower and upper boundaries within many subgroups of frequency combinations. While figure 3.2 plots all the three-tone results for both groups, it is easier to decipher the ratio boundaries by examining the distortion product classes individually. Figure 3.3 shows the results for distortion product classes in Group A, and Figure 3.4

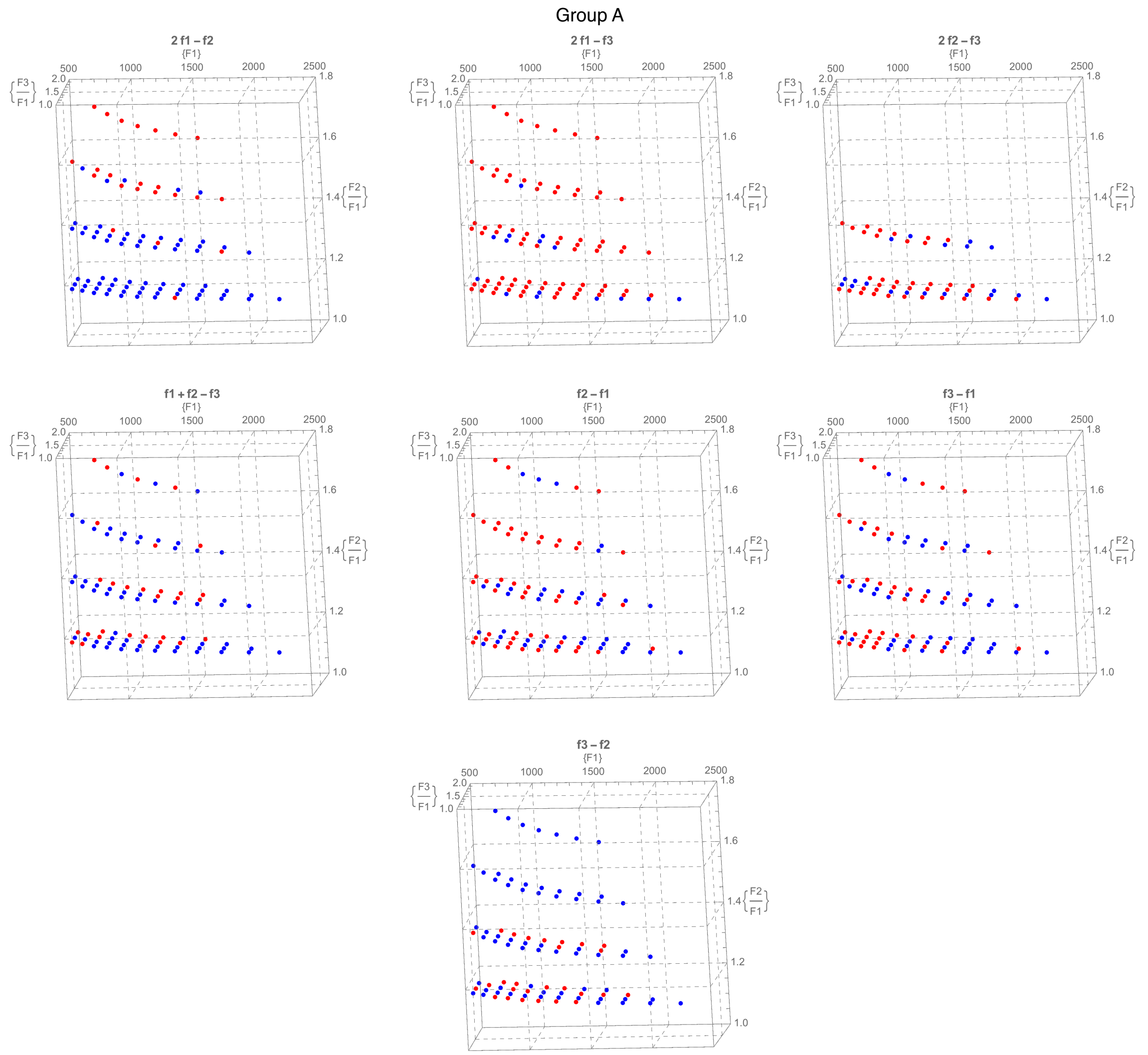


Figure 3.3. *All individual three-tone difference tone classes in Group A. The dots represent a difference tone above (blue) or below (red) threshold. The x-axis is the f_1 primary tone, the y-axis is the f_2/f_1 primary tone ratio, and the z-axis is the f_3/f_1 ratio.*

shows the results for distortion product classes in Group B. As previously discussed, the vertical tiers represent the nested subgroups within each block with a common f_1 primary. The y-axis is the f_2/f_1 ratio, and the z-axis height within each tier is the f_3/f_1 ratio. Ratio limits are found by evaluating the blue dots (above threshold) embedded by red dots (below threshold). In the cases where a boundary is not found, the recorded data is still useful for it reflects a range of frequencies that the population can reliably detect distortion products.

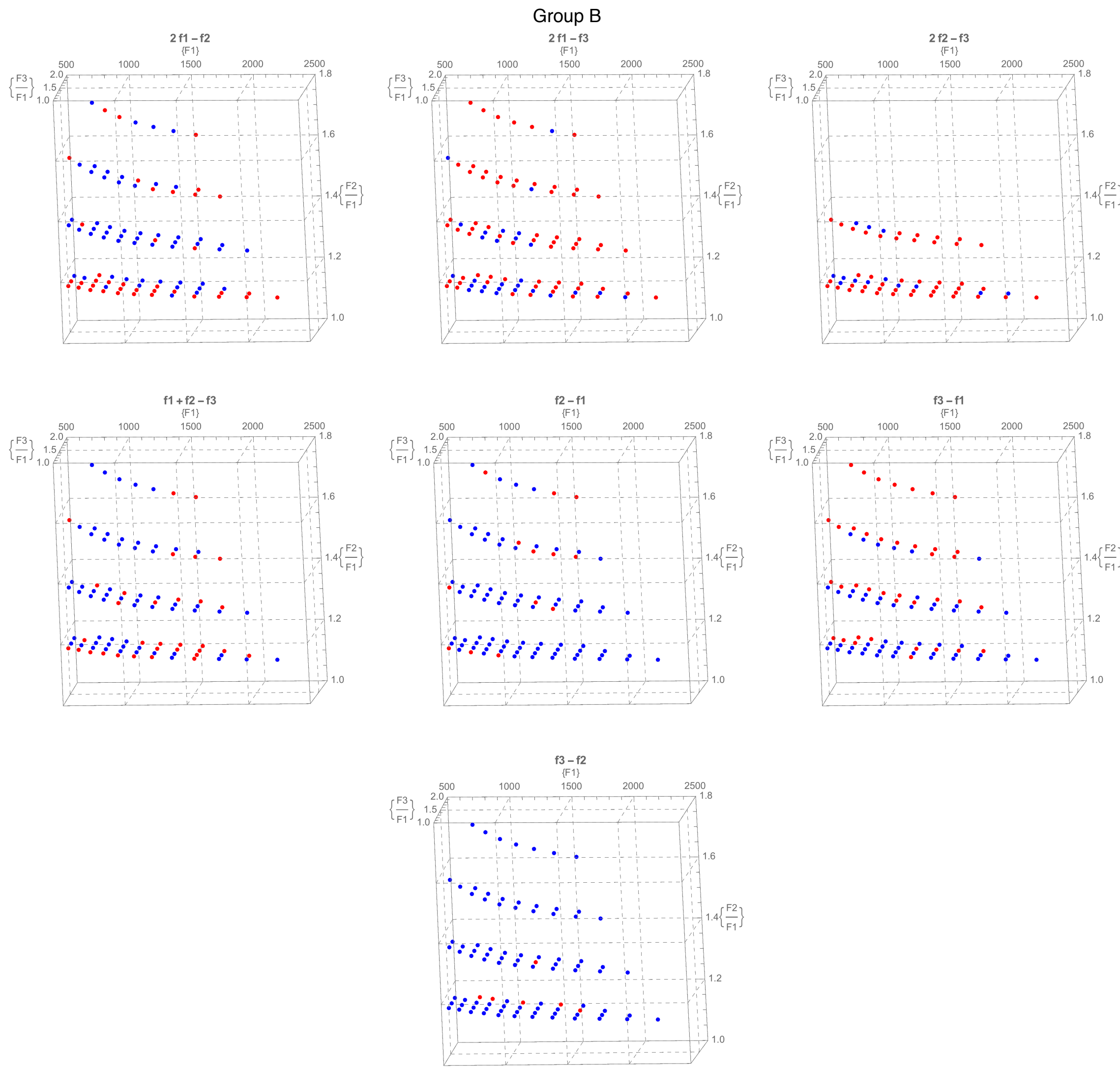


Figure 3.4. *All individual three-tone difference tone classes in Group B. The dots represent a difference tone above (blue) or below (red) threshold. The x-axis is the f_1 primary tone, the y-axis is the f_2/f_1 primary tone ratio, and the z-axis is the f_3/f_1 ratio.*

Figure 3.5 shows the mean and median number of three-tone difference tone classes above threshold within each f_1 primary tone block for Groups A and B combined. Comparison across groups is possible because there is not a significant difference for the metric of detection numbers ($t = -0.603$, $p > 0.55$). The average amount of detection across the frequency range is fairly consistent at around four difference tone classes per f_1 block. There is a slight dip in detection at 1196 Hz and a subsequent increase until the end of the frequency range. The increase above four detections in the final three f_1 blocks corresponds with the reduction of the number of wide ratio primaries in each block due to the ceiling effect. Since larger ratios typically produce less detection due to the ratio dependence of some difference tone classes, their absence produces higher average detection compared to f_1 blocks with the wider ratio ranges. Figure 3.6 shows the average number of classes above threshold for Group A and B separately. Group A exhibits a slight upward slope across the frequency range, although it is not consistent. The results for Group A show an average of approximately three to four classes above threshold for each f_1 block until 1539 Hz, and then a regular increase to seven cases by 2212 Hz. The results for Group B are generally higher until the upper f_1 frequency blocks. Overall, the average number of classes above threshold in each f_1 frequency block exhibits relatively even, but nonmonotonic behavior across the tested frequency region.

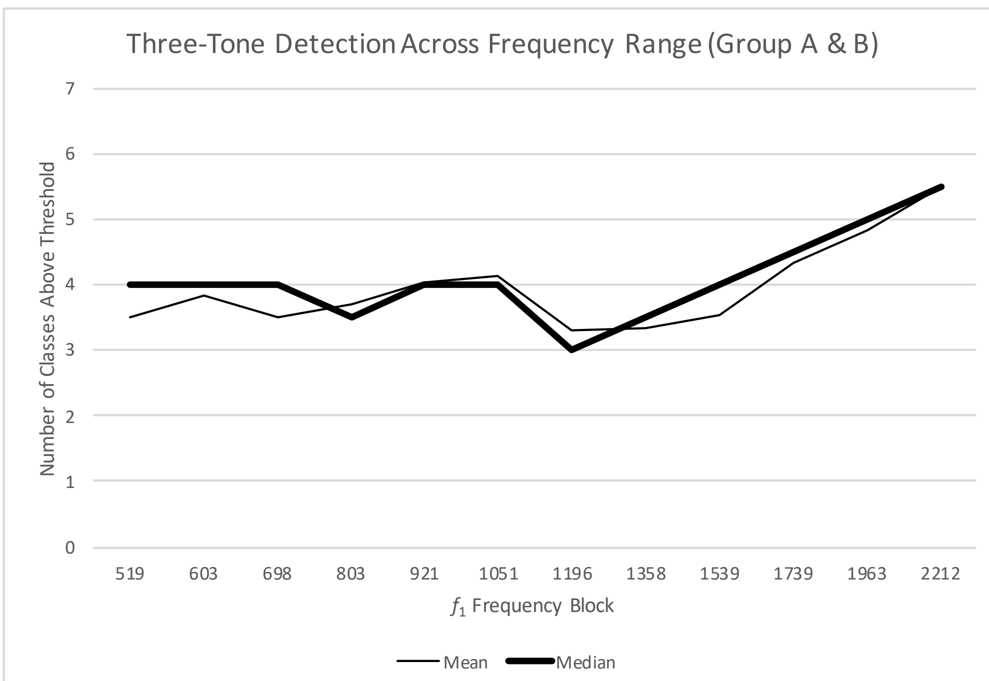


Figure 3.5. Mean and median number of three-tone difference tone classes above threshold across the frequency range for both groups. The x-axis is the given f_1 frequency block (in Hz), and the y-axis is the number of difference tone classes above threshold.

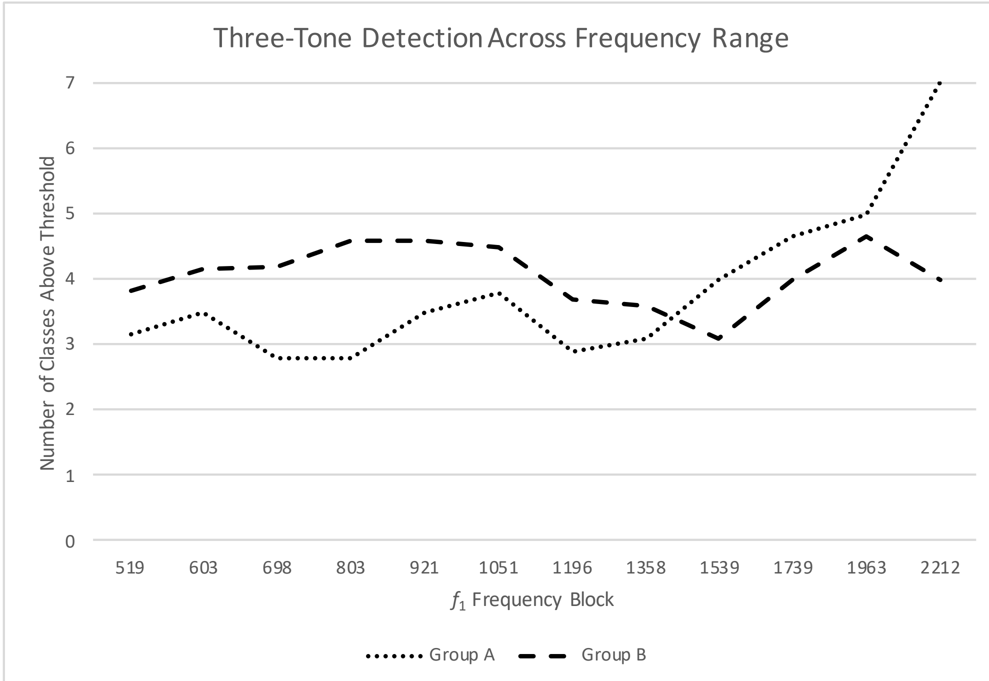


Figure 3.6. Average number of three-tone difference tone classes above threshold across the frequency range for each group individually. The x-axis is the given f_1 frequency block (in Hz), and the y-axis is the number of difference tone classes above threshold.

Similarly, Figure 3.7 shows the mean and median detection of difference tone classes across the f_1 frequency blocks for the two-tone difference tones in both groups. The groups can be combined for the metric does not show a significant difference (t = -1.406, p > 0.175). The median shows both classes of difference tones are above threshold across the frequency region, with only one exception at 1358 Hz. The average number of difference tone classes consistently varies just below the median values. Figure 3.8 provides the average number of two-tone difference tone classes above threshold for each group individually. The results continuously vary between 1.5 and 2, with only a single exception of a decrease to one case at 1196 Hz in Group A. Overall, the average number of audible two-tone difference tones is fairly even across the frequency region.

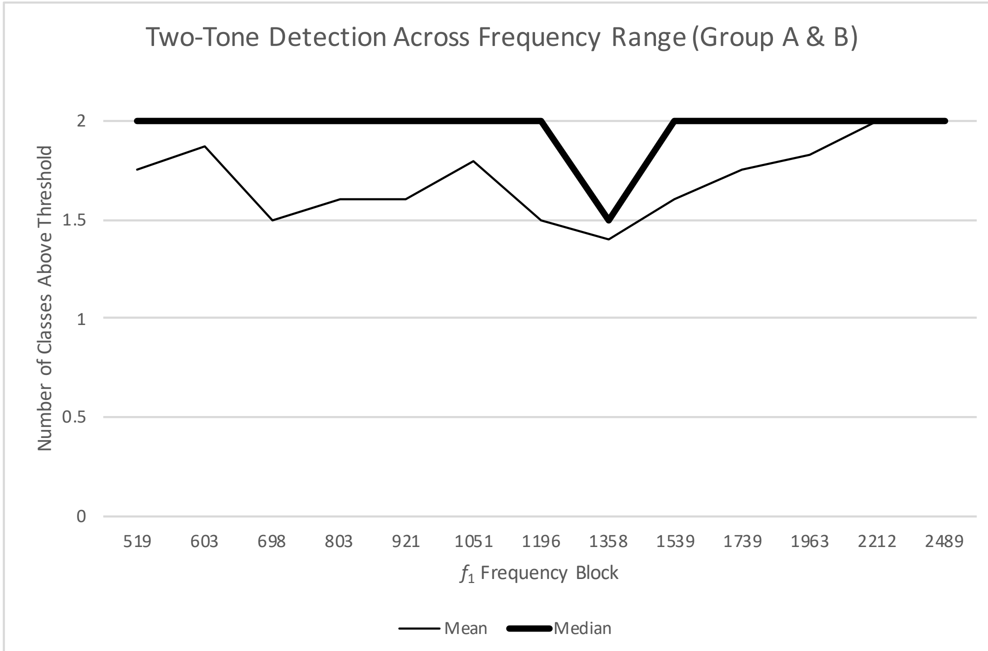


Figure 3.7. Mean and median number of two-tone difference tone classes above threshold across the frequency range for both groups. The x-axis is the given f_1 frequency block (in Hz), and the y-axis is the number of difference tone classes above threshold.

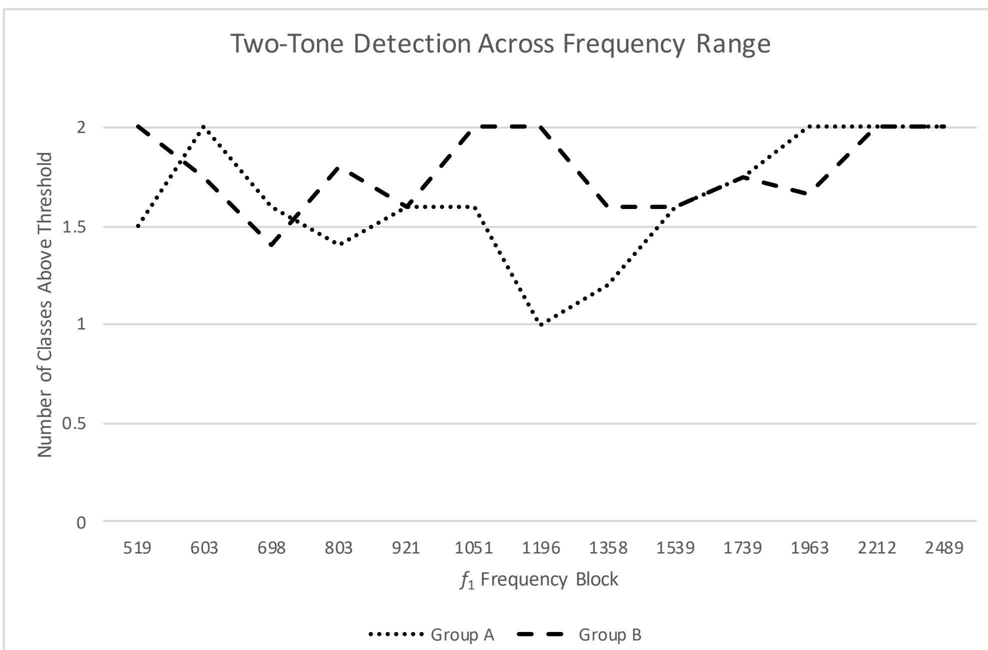


Figure 3.8. Average number of two-tone difference tone classes above threshold across the frequency range for each group individually. The x-axis is the given f_1 frequency block (in Hz), and the y-axis is the number of difference tone classes above threshold.

Next, let us examine the individual primary tone combinations within each f_1 frequency block. As we saw in the two-tone tables 3.5 & 3.7 and three-tone tables 3.6 & 3.8, the final columns show the number of difference tone classes above threshold for each stimulus. Table 3.11 (left) reports the same data for side-by-side comparison between both groups for two-tone difference tones. The relationship between ratio and detection is clear, with more difference tone classes above threshold at smaller f_2/f_1 ratios. The upper ratio boundaries are detected better in Group A. The number of cases with detection in both groups is large: 66% of the results in Group A and 77% in Group B show detection for both of the two possible difference tones. While there are no primary tone combinations that yield zero detection in both groups simultaneously, there are three sets of primaries that show zero detection in one group while the other group shows one case. All three of such cases occur at the maximum ratio possible for that particular f_1 primary tone, and they are in three consecutive f_1 frequency blocks (698 Hz, 803 Hz, and 921 Hz).

Table 3.11 (right) shows the side-by-side comparison showing the number of three-tone difference tone classes detected from specific primary tone combinations. Of the 93 conditions, there are three cases between the groups with all seven difference tones detected (the Group A condition of primary tones 1051 Hz, 1196 Hz, 1539 Hz, and 2212 Hz, 2489 Hz, 2798 Hz, and the Group B condition of 1051 Hz, 1196 Hz, 1739 Hz). There are no cases of all seven difference tones detected in both groups at the same primary tone condition, but there are three cases of six difference tones detected from the same primary tone condition (the 803 Hz, 1051 Hz, 1196 Hz primary tones, the 803 Hz, 1051 Hz, 1358 Hz primary tones, and the 1739 Hz, 1963 Hz, 2489 Hz primary tones) and 17 total cases with six difference tones detected. Also, there are 24 cases of the same number of difference tones detected in each group, of which 15 of these primary tones show detection of four or more difference tones. One notable case is the stimulus of 1196 Hz, 1539 Hz, 1963 Hz for it has zero detection in either group. This condition occurs in the middle of its subgroup. As the primaries both above and below this irregular condition show detection, perhaps the non-detection is the result of a reduction in difference tone amplitude due to nonmonotonic behavior.

f1	f2	Total A	Total B
519	603	2	2
519	698	2	2
519	803	1	2
519	921	1	2
603	698	2	1
603	803	2	2
603	921	2	2
603	1051	2	2
698	803	2	2
698	921	2	2
698	1051	2	1
698	1196	1	2
698	1358	1	0
803	921	2	2
803	1051	2	2
803	1196	1	2
803	1358	2	2
803	1539	0	1
921	1051	2	2
921	1196	2	2
921	1358	2	2
921	1539	2	1
921	1739	0	1
1051	1196	2	2
1051	1358	2	2
1051	1539	2	2
1051	1739	1	2
1051	1963	1	2
1196	1358	2	2
1196	1539	2	2
1196	1739	0	2
1196	1963	1	2
1196	2212	0	2
1358	1539	2	2
1358	1739	2	2
1358	1963	0	2
1358	2212	1	1
1358	2489	1	1
1539	1739	2	2
1539	1963	2	2
1539	2212	2	2
1539	2489	1	1
1539	2798	1	1
1739	1963	2	2
1739	2212	2	2
1739	2489	2	2
1739	2798	1	1
1963	2212	2	2
1963	2489	2	2
1963	2798	2	1
2212	2489	2	2
2212	2798	2	2
2489	2798	2	2

f1	f2	f3	Total A	Total B
519	603	698	2	2
519	603	803	4	4
519	603	921	5	6
519	698	803	2	4
519	698	921	4	4
519	803	921	2	3
603	698	803	3	3
603	698	921	5	5
603	698	1051	2	4
603	803	921	5	5
603	803	1051	3	4
603	921	1051	3	4
698	803	921	2	3
698	803	1051	4	5
698	803	1196	4	5
698	803	1358	2	2
698	921	1051	5	5
698	921	1196	5	5
698	921	1358	1	4
698	1051	1196	3	5
698	1051	1358	1	4
698	1196	1358	1	4
803	921	1051	3	4
803	921	1196	3	6
803	921	1358	2	5
803	921	1539	2	3
803	1051	1196	6	6
803	1051	1358	6	6
803	1051	1539	0	5
803	1196	1358	3	5
803	1196	1539	2	4
803	1358	1539	1	2
921	1051	1196	3	3
921	1051	1358	6	5
921	1051	1539	4	5
921	1051	1739	2	6
921	1196	1358	4	5
921	1196	1539	4	5
921	1196	1739	1	5
921	1358	1539	4	5
921	1358	1739	3	4
921	1539	1739	4	3
1051	1196	1358	3	3
1051	1196	1539	7	5
1051	1196	1739	4	7
1051	1196	1963	1	4
1051	1358	1539	4	5
1051	1358	1739	6	5
1051	1358	1963	4	5
1051	1539	1739	3	5
1051	1539	1963	3	2
1051	1739	1963	3	4
1196	1358	1539	3	2
1196	1358	1739	5	5
1196	1358	1963	3	3
1196	1358	2212	3	4
1196	1539	1739	4	6
1196	1539	1963	0	0
1196	1539	2212	3	5
1196	1739	1963	2	4
1196	1739	2212	3	4
1196	1963	2212	3	4
1358	1539	1739	1	6
1358	1539	1963	5	4
1358	1539	2212	2	3
1358	1539	2489	5	3
1358	1739	1963	4	4
1358	1739	2212	4	5
1358	1739	2489	3	3
1358	1963	2212	2	1
1358	1963	2489	4	4
1358	2212	2489	1	3
1539	1739	1963	5	3
1539	1739	2212	6	5
1539	1739	2489	4	2
1539	1739	2798	4	4
1539	1963	2212	5	4
1539	1963	2489	3	5
1539	1963	2798	3	3
1539	2212	2489	4	1
1539	2212	2798	4	3
1539	2489	2798	2	1
1739	1963	2212	6	4
1739	1963	2489	6	6
1739	1963	2798	5	3
1739	2212	2489	3	5
1739	2212	2798	6	3
1739	2489	2798	2	3
1963	2212	2489	6	5
1963	2212	2798	4	4
1963	2489	2798	5	5
2212	2489	2798	7	4

Table 3.11. A comparison between groups of the total number of two-tone (left) and three-tone (right) difference tones above threshold.

The $2f_1$-f_2 and f_2-f_1 difference tones classes are in both the two and three-tone conditions. While the two and three-tone data should be considered independently, it may be helpful to consider the differences between the two conditions. With only one exception, the addition of the third tone lowers the mean discrimination index and the percent of detection in both groups. For f_2-f_1, the mean discrimination index in Group A goes down from 0.4548 to 0.2199 and 0.5715 to 0.4586 in Group B. For the same difference tone, the percent of detection goes down from 89% to 43% in Group A, and 94% to 86% in Group B. The mean discrimination index for the $2f_1$-f_2 difference tone goes down from 0.4397 to 0.4127 in Group A and 0.4017 to 0.2615 in Group B. The percent of detection for $2f_1$-f_2 is the exception for it goes up in Group A from 68% to 75%, but it decreases in Group B with 81% for the two-tone condition to 59% in the three-tone condition. These results are likely due to the complexity of the three-tone stimulus and suppression effects (Humes, 1980b, 1983; Brown & Kemp, 1984; Harris *et al.*, 1992).

The results agree with past work on difference tone detection across primary tone ratio. For the two-tone stimuli, the closest comparison for primary tone frequency ratio is Plomp (1965, 1966). As shown in Figure 3.9, his investigation with a f_1=1000 Hz and equal level primaries of L_1=L_2=80 dB SPL approximately compares to the present study's stimulus data in the f_1=1051 Hz block (see tables 3.5 & 3.7). Between both groups, the f_2-f_1 difference tone detection is in the same range as Plomp, although the lowest and highest ratios in the current study do not have the resolution to show the exact boundary limits because all stimuli are above threshold. Plomp's lowest detectable ratio was ~1.04 and highest was between 1.8 and 1.9. In the present study, the lowest ratio is 1.38 and highest is 1.868. The Group B two-tone $2f_1$-f_2 difference tone in the current study also shows the same detection range as Plomp's $2f_1$-f_2 data. Plomp's lowest threshold ratio averaged around 1.1 and highest ratio averaged around 1.685. One of Plomp's subjects exhibited $2f_1$-f_2 detection through the upper ratio of 1.90, while the other three subjects exhibited the upper threshold limit of 1.52, 1.61, and 1.71. In the present study, the Group B $2f_1$-f_2 difference tone shows detection at all ratios in the block, resulting in a lower ratio boundary of 1.138 and an upper ratio boundary of 1.868. The current study's upper ratio boundary extends beyond the averaged values in Plomp, but this finding is not unexpected considering the additional 5 dB of stimulus level in the present study.

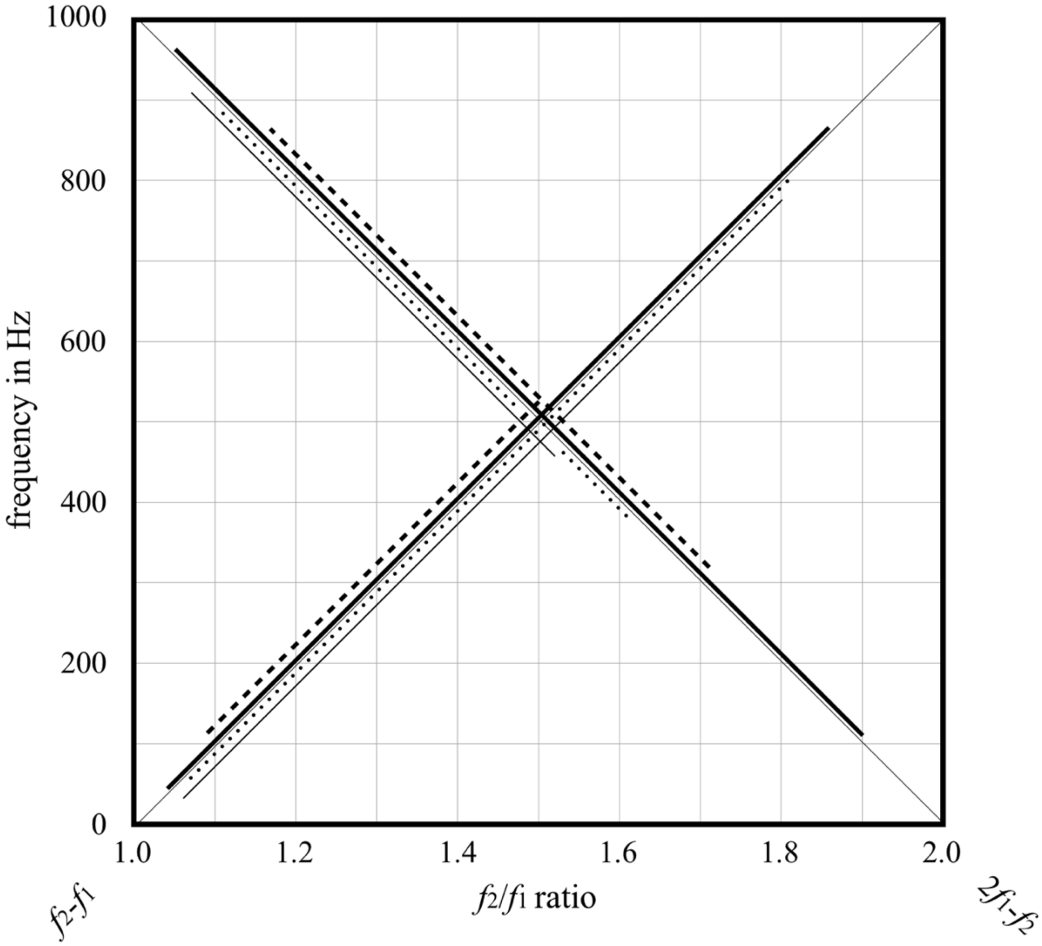


Figure 3.9. Replotted results from Plomp (1966) showing detection of difference tones across frequency range for four subjects. The thin lines connecting diagonally opposite corners are guides for potential f_2-f_1 and $2f_1$-f_2 difference tones, and the four subjects are denoted by regular, dotted, bold, and dashed lines. The f_1=1000 Hz and the primaries are equal amplitude at L_1=L_2=80 dB.

We cannot make a direct comparison between difference tone cancellation levels and detection because the difference tone levels were not measured, however it is possible to observe overall trends. Since there are only a few cases of non-detection in both groups of the two-tone f_2-f_1 difference tone class, the results show the relative ratio independence as addressed in the Introduction. That said, the range of ratios in the present study exceeds the range in the past work, and the few cases of non-detection generally occur at the highest ratios examined (i.e., 1.8, 1.9). The closest comparison would be Goldstein (1967), who measured up to the f_2/f_1 ratio 1.8, and saw a decline in QDT strength as the ratio increases. The results in the current study also show independence between f_2-f_1 detection and the f_1

frequency range, which is in accord with other work (Fastl & Zwicker, 2007). The two-tone $2f_1$-f_2 results show the established ratio dependence in the previously discussed CDT studies because the results exhibit consecutive non-detection at higher ratios. While the CDT ratio dependence occurs in both groups, the characteristic is stronger with the louder probe tone in Group A. This makes sense considering a lower amplitude probe tone will detect lower amplitude difference tones, and thus extend the range of distortion products above threshold.

Discussed in the previous chapter, the classical power series nonlinearity yields both the difference tone classes and their respective amplitude relationships. With three input frequencies, each with a numerical amplitude of 1, the QDT values (f_2-f_1, f_3-f_1, and f_3-f_2) each will have the amplitude of 1. The $2f_1$-f_2, $2f_1$-f_3, and $2f_2$-f_3 CDTs will have a reduced amplitude of 0.75 each, and the f_1+f_2-f_3 CDT will have an amplitude gain of 1.5. While the classical power series levels do not usually reflect difference tone cancellation levels, there are some similarities between the classical power series amplitude relationships and the current findings. Based on the results, there are more cases of the two-tone f_2-f_1 above the discrimination index than the two-tone $2f_1$-f_2 (see table 3.10). In Group B, there are more cases of QDTs that are generally above the discrimination index than CDTs. However, Group A exhibits a more even mix between the all the possible CDTs and QDTs, suggesting the classical power series amplitude hierarchy is not in place. Also, according to the classical power series nonlinearity, the CDT f_1+f_2-f_3 should be the loudest, but that particular difference tone class is in the middle rank of difference tones above threshold. Past work investigating both the f_2-f_1 and $2f_1$-f_2 in the same experiment often showed the $2f_1$-f_2 difference tone has a higher amplitude than the f_2-f_1 difference tone (Zwicker, 1979a; Oxenham *et al.,* 2009), or at least is roughly equal in amplitude (Hall, 1972a). Overall, the ranking of difference tones above threshold in the current study does not match the hierarchy found in the classical power series nonlinearity, which is in line with past experimental findings.

A few details should be considered when interpreting the results. Because the experimental design was a straightforward detection task that did not involve an adjustable cancellation tone, it is important to consider the possibility that some beats were generated by means other than the fixed probe tone interacting with the intended difference tone.

While the selected primary tones were ERB center frequencies, and the difference tone classes were selected for producing difference tones below the lowest primary, some of resulting difference tones were generated in the same ERB band as a primary tone. In these cases, a traditional beat or roughness may be perceived and misinterpreted as intended beats, or the difference tone may be masked by the primary tone. The difference tone classes affected were the $2f_2$-f_3 with 16 cases (or 17 if counting the probe tone at 3 Hz above the DT), and f_3-f_1 with four cases. For example, the stimulus consisting of 519 Hz, 603 Hz, and 698 Hz produces a $2f_2$-f_3 difference tone at 508 Hz, which is well within the calculated 81 Hz ERB bandwidth of the f_1 519 Hz primary tone. However, with only one exception, discrimination for all cases was below threshold in both groups, which suggests the difference tones are masked by the f_1 primary tone. The only case with detection above threshold was the stimulus 2212 Hz, 2489 Hz, 2798 Hz in Group A, which yielded a $2f_2$-f_3 difference tone of 2180 Hz (the ERB is 263 Hz, or 132 Hz below f_1). This was also the highest tested stimulus frequency combination.

Other phenomena to consider are beats that arise due to mistuned consonances, phase effects between the stimulus components, and internal interactions between the distortion products (Plomp, 1967, 1976; Hall, 1972b; Buunen *et al.*, 1974; Buunen & Bilsen, 1974; Buunen, 1975; Feeney, 1997). While the stimulus components were not harmonically related because they were ERB center frequencies, some primary tone relationships could have been within a range that qualifies as mistuned consonances. The phase relationship between the primaries could lead to the detection of beats due to the auditory system's sensitivity to the envelope of the combined signal. Unintended beating could also result from interactions between two or more audible distortion products. For example, the stimulus consisting of 519 Hz, 803 Hz, and 921 Hz produced a f_3-f_1 difference tone at 402 Hz and a f_1+f_2-f_3 difference tone at 401 Hz, a 1 Hz difference between difference tones. To partially combat these effects, the participants were trained to answer "beat" if they heard approximately six pulsations within the two-second trial, although counting was not encouraged. The beating was not likely the result of the 3 Hz probe tone if it was much faster or slower, and the participant responded accordingly. Also, because the discrimination index was the difference between the hit rate and false alarm rate, the result compensated for some of these potential errors. All issues considered, the specific difference tones and primary tones are included in tables 3.5-3.8, and the results can be interpreted with these considerations in mind.

Three technical issues occurred during testing. About half of the subjects reported a brief, irregular clicking sound that lasted for approximately 5-15 seconds. The issue is suspected to be a result of the audio routing utility Soundflower. After it was a known issue, the possibility of its occurrence was mentioned during training and the participant was encouraged to answer to the best of their abilities. Another issue involved one participant accidentally aborting the experiment after a period of 45 minutes of testing. The partial session was discarded, and the subject started again. Finally, a PsychoPy crash during testing led to one participant restarting the experiment after nearly an hour of previous testing. However, given the limited duration and randomness of the clicks, and random stimuli presentation in PsychoPy, these technical issues were minor and likely did not influence the overall results.

The results form a large dataset of up to nine difference tones evoked from two and three-primary tones across a wide frequency spectrum and multiple primary ratios. With a focus on detection rather than amplitude measurement, the experimental design allowed for many primary tone combinations to be examined in a single session. As a result, an increased primary tone frequency range and number of primary tones were employed. While past work typically examined a smaller number of difference tone classes, primary tones, and primary tone ratios, the current method minimizes the need for concatenating data from multiple smaller experiments with potentially difference equipment, methods, and facilities. The results can be of use in future perceptual and psychoacoustical studies, as well as music perception, instrument design, and creative work. The application of the results in technology and creative contexts will be discussed in chapters to follow.

3.5 SUMMARY AND CONCLUSION

The purpose of this study was to investigate the perception of cubic and quadratic difference tones resulting from two and three-tone stimulus across a wide frequency and primary tone ratio range. Employing an experimental design measuring detection allowed for the examination of nine classes of difference tones produced from a primary tone f_1 range between 519 Hz and 2489 Hz, and ratios between 1.12 and 1.95. The main questions were to determine which difference tones classes are the most detectable, how many difference tones can be detected from a single primary tone combination, how much does detection vary across the frequency range, and what are the ratio boundaries for each difference tone class.

Perceivable difference tones were found within all nine classes examined. General detectability for each difference tone class was determined by calculating the mean discrimination index, as well as evaluating the percent of individual cases above threshold within each class. Both the two-tone f_2-f_1 and $2f_1$-f_2 difference tone classes were found to have mean discrimination indices above threshold, and the three-tone $2f_1$-f_2, f_3-f_2, f_1+f_2-f_3, and f_3-f_1 difference tones were above threshold in both groups. Among all nine classes of difference tones, the two-tone f_2-f_1 QDT had the highest magnitude discrimination index and the highest percentage of cases above threshold in both groups. The three-tone difference tone with highest magnitude discrimination index was $2f_1$-f_2 in Group A, and f_3-f_2 in Group B. The same two difference tone classes also had the highest percentage above threshold in their respective groups.

Both the f_2-f_1 and $2f_1$-f_2 difference tones were simultaneously perceived in the majority of the two-tone stimuli (66% in Group A, 77% in Group B), and as many as seven difference tones were simultaneously detected from select three-tone primaries. The two-tone difference tones exhibited fairly consistent yet nonmonotonic behavior across the frequency spectrum, with an average of 1.6 and 1.8 difference tone classes detected for Group A and B respectively. The three-tone difference tones were also consistently detected in nonmonotonic behavior across the spectrum, yielding an average of 3.9 and 4.1 difference tone classes in Group A and B. There was an increase in the amount of three-tone difference tone classes detected in the highest f_1 frequency blocks in Group A. The increase may be the result of the decreasing number of stimuli tested within the highest f_1 primary blocks due to the ceiling effect, which also corresponds to a reduction of higher ratio primary combinations.

Primary tone ratio boundaries were detected in all nine classes of difference tones to varying degrees. The upper ratio boundary was mostly visible for the two-tone $2f_1$-f_2

CDT, which is consistent with the monotonic amplitude decrease with increasing ratio when L_1=L_2 in past work. The lower limit generally was not found, likely due to the limited ratio resolution from using ERB center frequencies as primary tones. The two-tone f_2-f_1 QDT produced high detection in both groups, which resulted in fewer ratio boundary limits. These findings point towards the relative independence between even-order difference tones and primary tone ratio as seen in literature. The relatively few cases of non-detection generally occurred at higher ratios, which is consistent with the decrease in QDT amplitude at higher ratios in work by Goldstein (1967), and Hall (1972a).

Notable ratio boundaries and other details involving the three-tone difference tone classes are as follows. Like its two-tone counterpart, the $2f_1$-f_2 also showed a dependence on primary tone ratio with less detection at higher primary tone frequency separations. As a result, the maximum ratio boundary is clear in many of the f_1 primary tone blocks. More minimum ratio boundaries were found in Group B than Group A. The $2f_1$-f_3 and $2f_2$-f_3 CDTs had a mean discrimination index below threshold, but sparse detection shows perceivable difference tones with clear ratio boundaries. The f_1+f_2-f_3 CDT showed a ratio dependence because many of the subgroups exhibited maximum ratio boundaries, particularly in Group A. These results are consistent with the investigation by Weber and Mellert (1975), which found a nonmonotonic amplitude decrease as the f_3 primary increased. The three-tone f_2-f_1 QDT was the only class with a mean discrimination index above threshold in one group and below threshold in the other. While below threshold in Group A, the detection of individual f_2-f_1 difference tones within the class was at least twice as high as the other below-threshold classes in the same group. A reasonable explanation is the mean discrimination index (0.2199) is only marginally below threshold (0.2294). The f_3-f_1 QDT contained a large amount of irregular non-detection within the subgroups, and of the classes with an above-threshold mean discrimination index, it was the difference tone class with the lowest percent of detection. Finally, the f_3-f_2 difference tone class exhibited strong detection, with only a few cases of non-detection (particularly in Group B). As a result, neither group contained f_1 block maximum ratio limits. This characteristic is consistent with the ratio independence found in past work investigating two-tone quadratic difference tones. In the cases where a particular ratio boundary was not found, the resulting data still shows which primary tone combinations yield detectable difference tones.

When comparing the two and three-tone versions of the same difference tone combination, the addition of a third primary tone decreased both the mean discrimination index, and the percentage of detection within the difference tone class. The only exception to this finding was the $2f_1$-f_2 in Group A, which had a higher percentage of detection in the three-tone version (it increased from 68% to 75%).

It is possible for some difference tones to be masked if the probe tone level is too loud. Given that the difference tone classes each exhibit unique amplitudes, and the levels vary between subjects, a universal fixed probe tone level is unlikely. Oxenham *et al.* (2009) reported average probe tone levels that produced salient beats were L_1=L_2-20 dB (CDT) and L_1=L_2-40 dB (QDT). In the results of the present study, there were more difference tones detected with the lower amplitude probe level (L_p=L_1-30 dB), so it would be of interest to see the results with a probe level of L_1-40 dB in future work. At what level does the probe tone become insufficient? Is it possible to detect beats with a probe tone, but not perceive the difference tone without the aid of the probe tone? These are questions to consider in future work.

The role of phase in the current study is another topic of potential future research. The phase relationship between three-component primaries can change the level of the resulting difference tones (Buunen *et al.*, 1974; Buunen, 1975; Pressnitzer & Patterson, 2001; Oxenham *et al.*, 2009; Nuttal *et al.*, 2018). For example, Oxenham *et al.* found a small but significant trend that the estimated f_2-f_1 QDT and $2f_1$-f_2 CDT amplitudes were stronger when the starting phases of all three components were the same, as opposed to an alternate phase relationship with the middle component shifted 90°. It would be of interest to see the effect of a phase-shifted middle component for comparison with the results of the present experiment.

The experimental task in the present study was designed to test the detection of distortion products from a large number of stimuli in a single session. Other research interests in future work include examining higher order difference tones (quartic, quintic, etc.), as well as four or more primary tone combinations. It is possible to examine these conditions with either no or minor adjustments to the developed paradigm.

CHAPTER 4
The Ear Tone Toolbox

4.1 INTRODUCTION

This chapter details the ongoing development of digital instruments designed to evoke auditory distortion products with precision and clarity. The software was initially created for single use purposes in specific compositions (to be discussed in Chapter 5), but soon evolved into generalized and simplified unit generators. A number of recent papers addressed synthesis techniques utilizing difference tones (Haworth, 2011, 2012; Kendall *et al.*, 2012, 2014), however, no widely released music software has been available. The work described in this chapter serves to fill this gap, while also operating as an accessible educational tool for understanding the underlying principles involved. The instruments were built from examining particular difference tones, or relationship between tones, and finding the appropriate equations algebraically. Most of the instruments described in this chapter were developed prior to the perceptual study from Chapter 3, however they can be used in parallel with the findings to evoke specific difference tones that are known to be perceptually reliable. The following text describing the software is built upon and expanded from portions of a paper presented at the 2016 International Computer Music Conference in Utrecht, Netherlands (Chechile, 2016a).

4.2 THE EAR TONE TOOLBOX

The Ear Tone Toolbox (ETT) is a collection of instruments for the production of auditory distortion product synthesis. In a variety of techniques, the toolbox generates the necessary acoustic primary tone combinations for evoking difference tones in the inner ear. The open-source software is developed in the FAUST (Functional AUdio STream) programming language for real-time audio signal processing (Orlarey *et al.*, 2009), and as a result, the ETT can easily compile to many architectures and formats. In its current state, the ETT offers external objects for Max/MSP and Pure Data, VST & VSTi plugins for standalone use in digital audio workstations, and patches for use with hardware and programmable modular synthesizers (Teensy, OWL eurorack module/pedal). In addition to the source code, an executable version of the Ear Tone Toolbox is available for Macintosh and Linux platforms.

The toolbox consists of five main unit generators that allow the user to specify various combinations of distortion products and acoustic primary tones. Most of the instruments have been constrained to function within primary tone ratios that are known to generate distortion products, but many are also available in unlimited versions. In the latter case, it is up to the user to play within an effective range if the desired effect is to evoke difference tones.

Distortion Product Focus with DiffTone

DiffTone generator allows the direct synthesis of user defined auditory distortion products. By specifying the desired f_2-f_1 QDT (f_Q) and $2f_1$-f_2 CDT (f_C) frequencies, the instrument produces the acoustic primary tones f_1 and f_2 for evoking the distortion products with the equations $f_1 = f_Q+f_C$ and $f_2 = 2f_Q+f_C$. For example, if a 500 Hz QDT and an 1100 Hz CDT were input, the object would generate two sine tones at f_1=1600 Hz and f_2=2100 Hz. In reverse, we see the two primary tones create the desired combination tones with our original equation for the QDT as 2100-1600 = 500 Hz and the CDT as 2*1600-2100 = 1100 Hz.

The *DiffTone* generator, along with the other instruments in the toolbox, contains an optional guide tone that can be used for testing audibility of the difference tone, demonstrating the effect by providing a tone to listen for, and other educational purposes such as ear training, however the guide tone should remain off during normal use as it will mask the internal ear tones. The first and second outlets of the Max object provide the respective f_1 and f_2 sinusoids and the third and fourth outlets provide the sinusoid guide tones at the f_2-f_1 QDT and $2f_1$-f_2 CDT frequencies. Figure 4.1 shows the *DiffTone~* help file for Max/MSP. Like all ETT Max objects, the input parameters are specified using the prepend object. For example, to evoke a 500 Hz QDT and 1100 Hz CDT, the output of each respective number box will connect to a prepend QDT object and a prepend CDT object, and the outputs will connect to the *DiffTone~* object. After the user makes these connections, the individual frequencies in the number boxes can be changed at will for immediate playback.

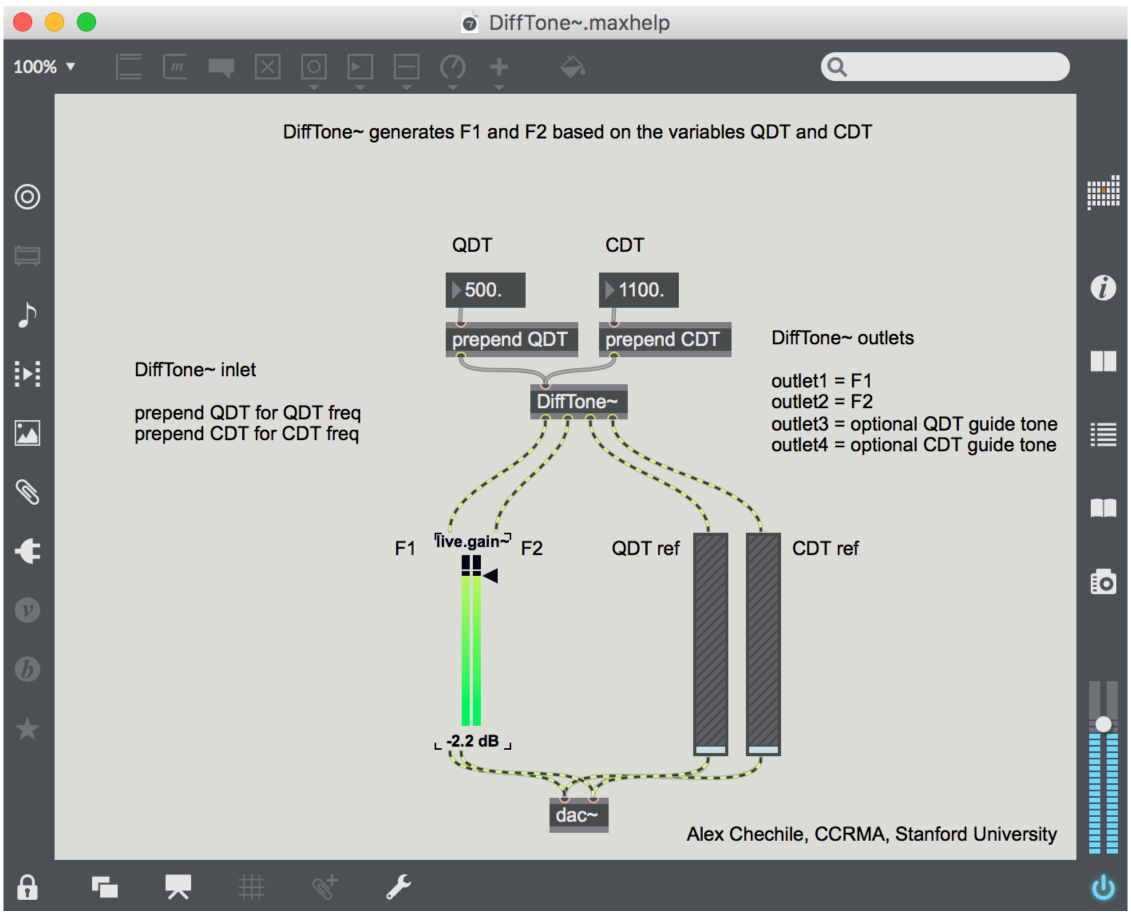


Figure 4.1. The Max help file for DiffTone~ provides an overview of the input parameters and output signals of the external object.

Primary Tone Focus with *f1ratio*

The *f1ratio* unit generator is the opposite of *DiffTone~* for it is focused on the primary tones by allowing direct control over the f_1 frequency and the ratio between the second primary. Distortion products occur as a physiological response to the acoustic primary tones f_1 and f_2, and the audibility of the resulting difference tone is in many cases dependent on the ratio between the primaries. The cubic difference tones, for example, are particularly sensitive to the frequency separation between the primaries. In the recording of distortion product otoacoustic emissions, the f_2/f_1 ratio of around 1.22 is typical for it has been shown to evoke maximal $2f_1$-f_2 emission levels (Harris, *et al*., 1989; Gaskil & Brown, 1990; Nielsen, *et al.*, 1993). Hence, *f1ratio* allows for direct dynamic

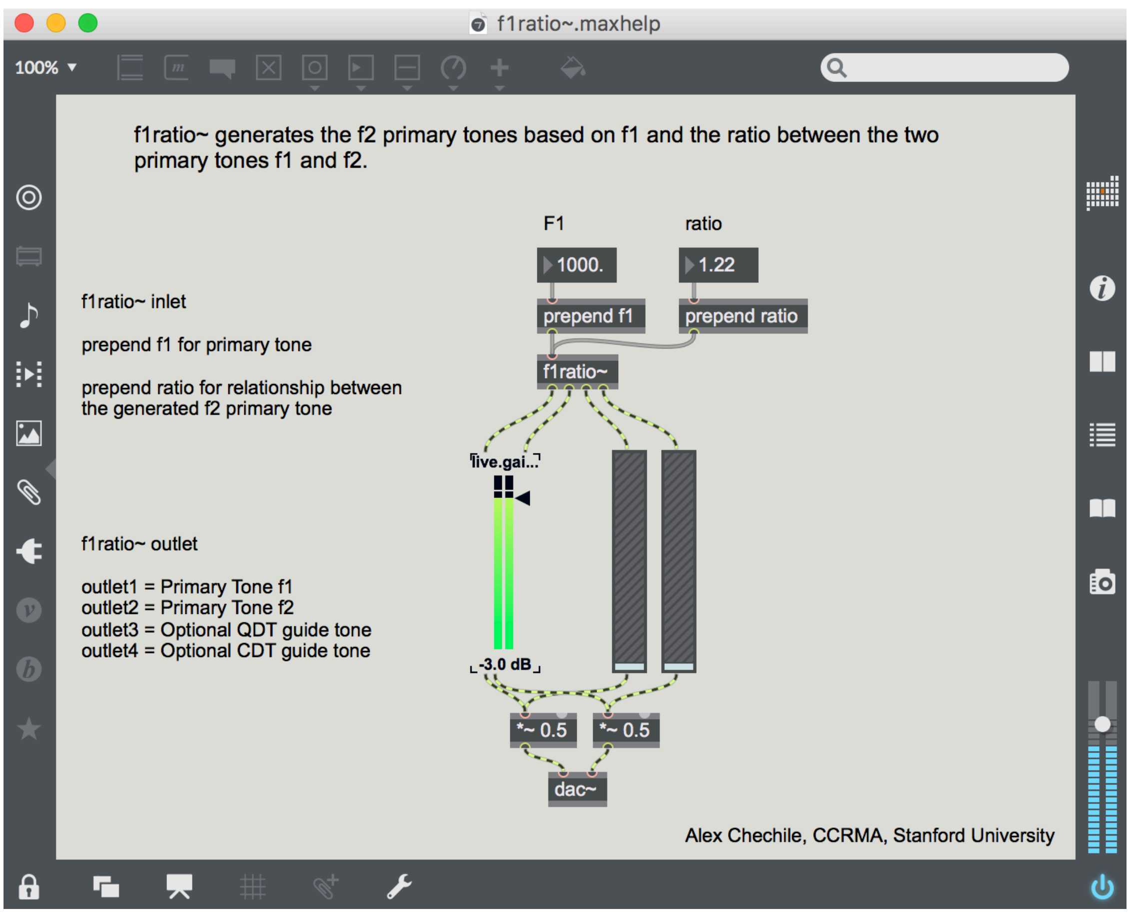


Figure 4.2. The Max help file for f1ratio~ illustrates the input and output parameters of the external object.

control of the ratio between stimulus tones. In the instrument, the f_1 is specified by the user, and the f_2 is calculated with the equation $f_2=f_1*r$, with r representing the ratio. With *f1ratio~* Max object, the input parameters are specified with prepend f_1 and prepend ratio, and it produces the f_1 and f_2 sinusoids from the first two outlets, and the optional guide tones for the QDT and CDT from the third and fourth outlets. The Max/MSP help file for *f1ratio~* is shown in Figure 4.2.

Simultaneous Distortion Product and Primary Tone Focus with *f1half* and *f2half*

For applications requiring specific control over both an acoustic component and a distortion product, the following two objects are available. The unit generators *f1half* and *f2half* (in which the term *half* refers to half primary tone control, and half difference tone

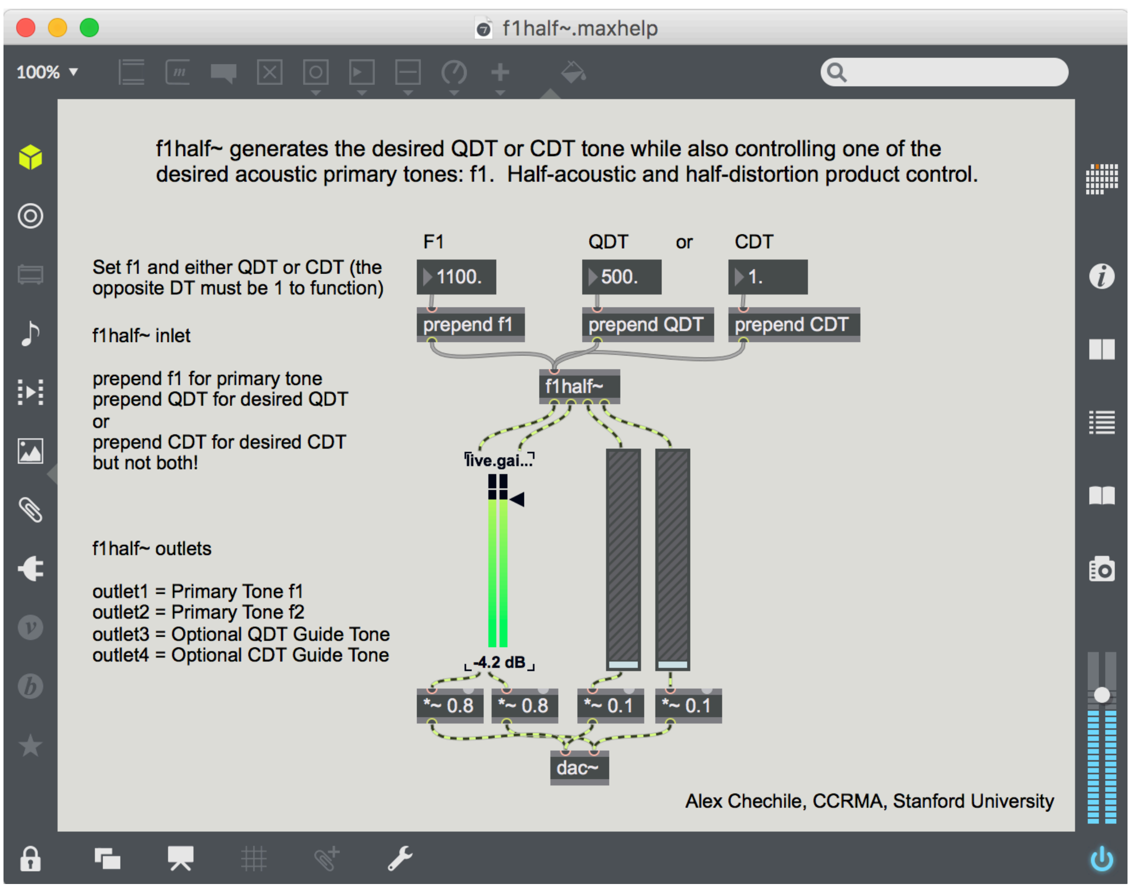


Figure 4.3. The Max help file for f1half~, which is similar to f2half~. Note one of the QDT or CDT values in this patch must be set to "1" for the object to calculate the second primary tone.

control) allow the user to specify one of the acoustic primary tones (the f_1 in *f1half* and f_2 in *f2half*) and either the QDT or the CDT. When the QDT is specified, *f1half* calculates the f_2 with the equation $f_2=f_Q+f_1$, and when the CDT is specified, *f1half* calculates the f_2 with the equation $f_2=(2f_1)-f_C$. Similarly, *f2fhalf* calculates the f_1 frequency with the equation $f_1=f_2-f_Q$ when the QDT is specified, and $f_1=(f_C+f_2)/2$ when the CDT is specified. It is important to note that both objects require only one specified combination tone while the other must be set to "1" (yet the uncalculated second distortion product will still be produced in the ear). The Max help file for *f1half~* is shown in Figure 4.3. The first two outlets provide the sinusoid primary tones f_1 and f_2, and outlets three and four provide optional guide tones for the respective QDT and CDT frequencies.

Three-Tone Focus with *DiffTone_3t*

Up to this point, all of the difference tones discussed in this chapter are the result of two primary tones. The *DiffTone_3t* unit generator utilizes three primary tones to allow direct control over the $f_1+f_2-f_3$ CDT. The instrument functions similarly to *DiffTone* for the QDT and CDT are also definable. *DiffTone_3t* generates the appropriate three primaries for evoking all three difference tones using the following formulas: $f_1=f_Q+f_C$, $f_2=f_1+f_Q$, and $f_3=3f_Q+2f_C-f_{CDT2}$ (or more simply, $f_3=f_1+f_2-f_{CDT2}$), where f_{CDT2} is the desired $f_1+f_2-f_3$ frequency. Figure 4.4 shows the max help file for *DiffTone_3t~,* with the QDT input specified by "prepend QDT," the CDT input by "prepend CDT," and the $f_1+f_2-f_3$ CDT input specified by "prepend f1+f2-f3." The first three outputs are the respective f_1, f_2, and f_3 primaries, and like the other objects in the collection, the fourth, fifth, and sixth outputs provide optional guide tones for the QDT, CDT, and $f_1+f_2-f_3$ CDT.

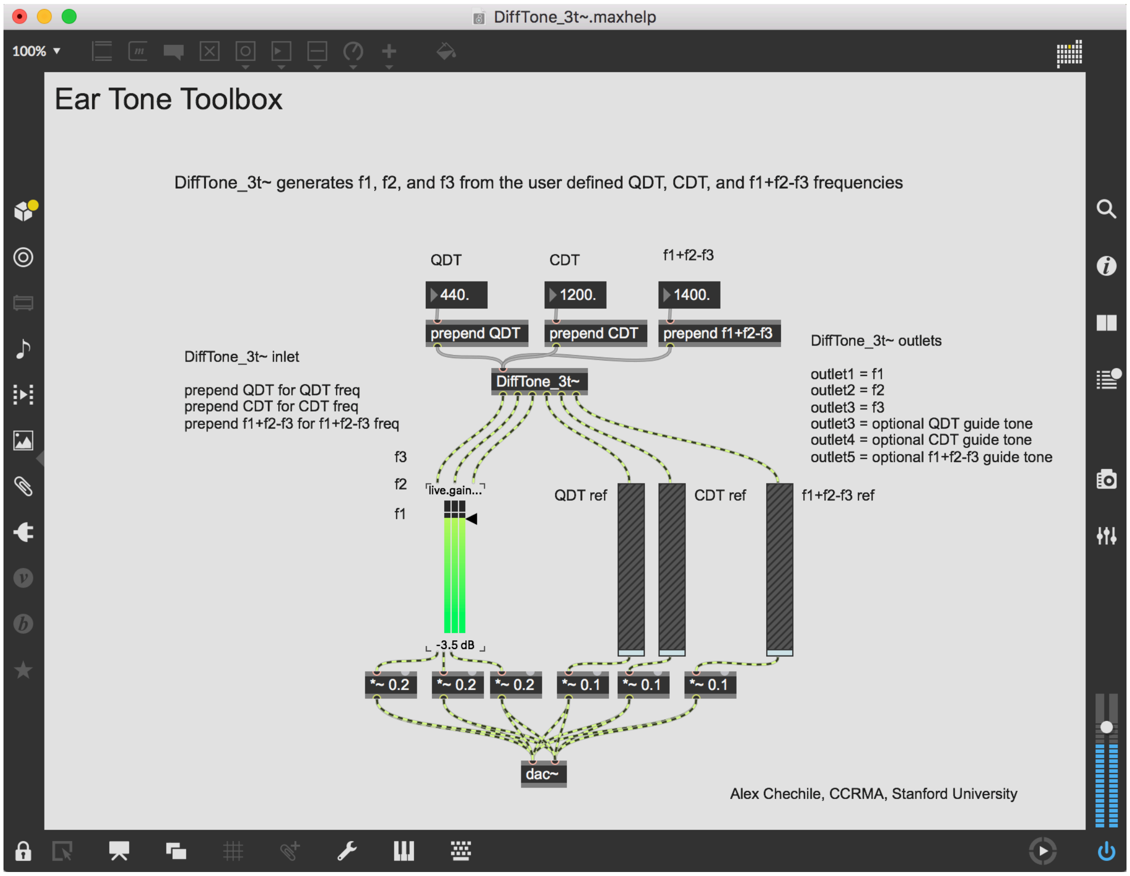


Figure 4.4. The Max help file for the three-tone DiffTone_3t~ unit generator, showing the various input and output parameters.

Distortion Product Spectrum with *DPSpecS* and *DPSpec*

As discussed in Chapter 1, Pressnitzer & Patterson (2001) found the magnitude of the QDT increased with an increasing number of consecutive harmonic primaries (2-17 components), and they found a distortion product frequency spectrum above hearing threshold. The even spacing between each sequential harmonic is equal to the f_2-f_1 QDT, and the difference between alternating harmonics is equivalent to the second harmonic of the QDT, etc. In an example with 1.5 kHz, 1.6 kHz, 1.7 kHz and 1.8 kHz primaries, the f_2-f_1 is 100 Hz (1.6-1.5 kHz), and the harmonic spectrum of distortion products would be produced: 200 Hz from f_3-f_1 (1.7-1.5 kHz), and 300 Hz via the f_4-f_1 (1.8-1.5 kHz) combinations. The magnitude of the f_0 QDT (f_2-f_1, f_3-f_2, f_4-f_3, etc.) increases regularly by ~3 dB per pair doubling of primary components, suggesting each pair contributes to the measured amplitude (Pressnitzer & Patterson, 2001).

The *DPSpecS* unit generator creates a distortion product spectrum following the Pressnitzer and Patterson study, as well as Haworth (2011). The user specifies the f_1 acoustic primary tone as well as the distortion product fundamental f_0, and the instrument produces a spectrum of sinusoids spaced by the value of the f_0. For example, if the user specifies a f_1 of 1000 Hz and a 100 Hz QDT f_0, the object will output twelve sinusoids in stereo (alternating six tones from the first outlet and the other six from the second outlet) spaced by 100 Hz. The third through sixth outlets of the Max object provide optional guide tones for the distortion product fundamental and the next three harmonics. A multichannel version of the instrument, *DPSpec* is also included in the ETT and provides individual outlets for each sinusoid in the spectrum. The spectrum of primary tones is calculated by $f_n=(n\text{-}1)f_0+f_1$ where f_0 is the distortion product fundamental and $n \geq 2, 3 \ldots 12$.

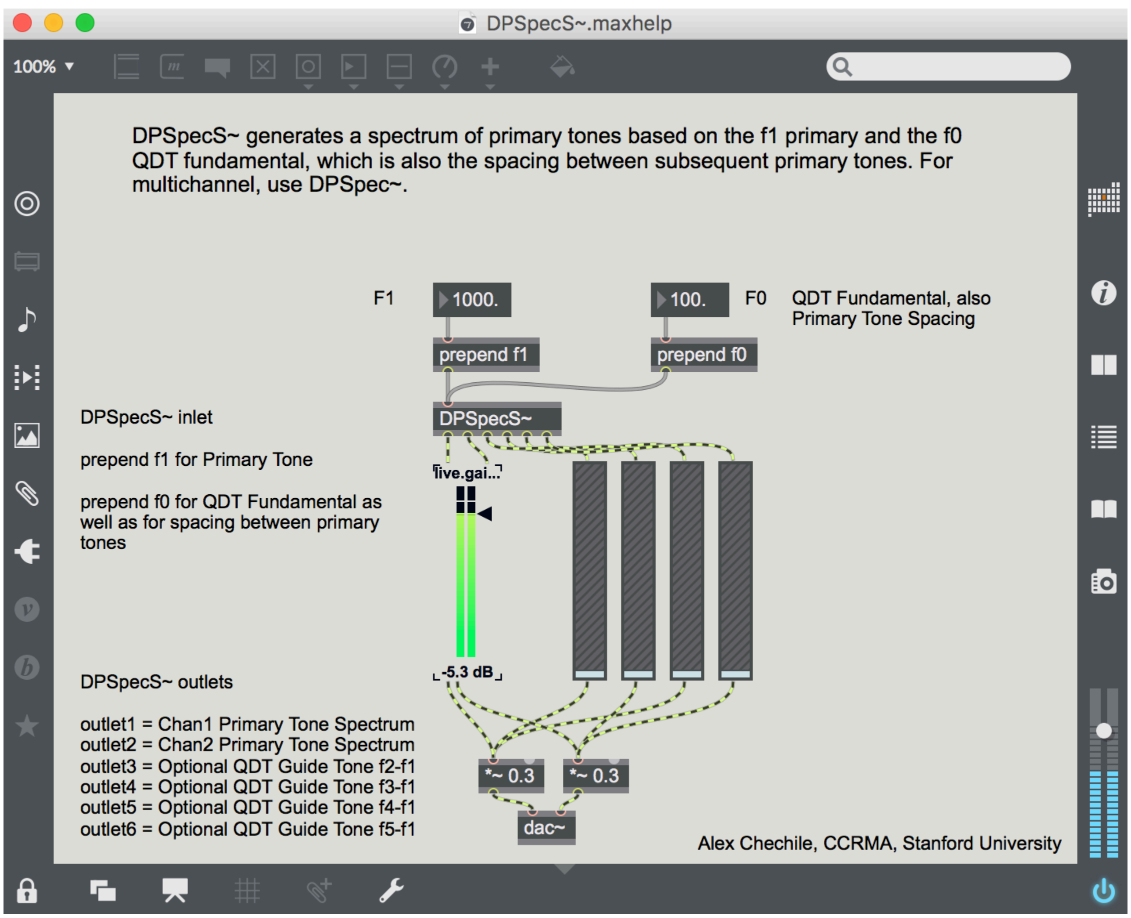


Figure 4.5. The Max help file for DPSpecS~ illustrating the input parameters and the output signals.

Continuing our example, the QDT between f_2 and f_1 equals the 100 Hz DP fundamental f_0, as does the difference tone between f_3 and f_2, etc. The CDT between f_1 and f_2 is also generated at 900 Hz. Distortion products between the harmonics are also produced, although at a lower amplitude. For example, a 200 Hz QDT is produced between f_3 and f_1, and between f_4 and f_2, etc. Figure 4.5 shows *DPSpecS~* Max help file.

4.3 IMPLEMENTAION

The Ear Tone Toolbox Max/MSP and Pure Data objects allow for easy implementation in programming environments. The instruments are also available as VST and VSTi plugins for direct use in digital audio workstations such as Ableton Live (Figure 4.6). The VST plugins are not signal processing effects, but rather versions of the instruments that can be controlled either manually or through automation. Any incoming audio will be passed through the VST in the signal chain. The VSTi plugins allow for MIDI control, so the user can play with any instrument controller. The VSTi plugins are available in both monophonic and polyphonic versions.

In addition to plugins and unit generators for software sound synthesis, the ETT is available in dedicated versions for hardware synthesizers. Built by Rebel Technology,

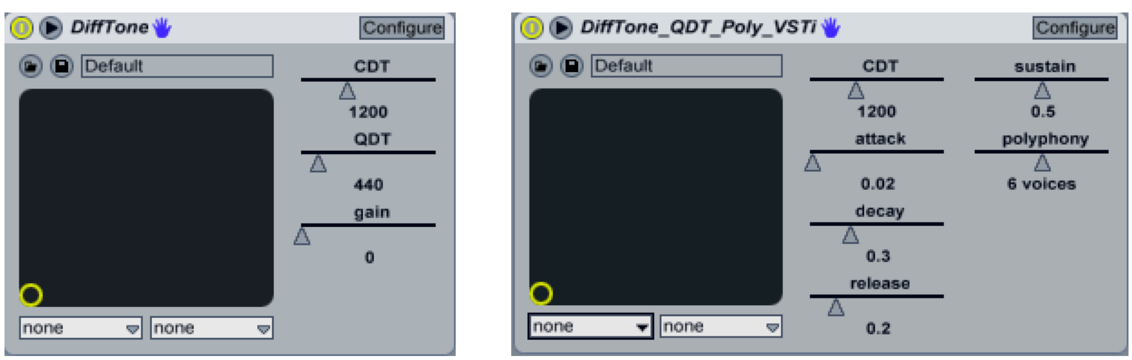


Figure 4.6. Standalone versions of DiffTone as a VST (left) and VSTi instrument (right) for use in Ableton Live, or other compatible digital audio workstations.

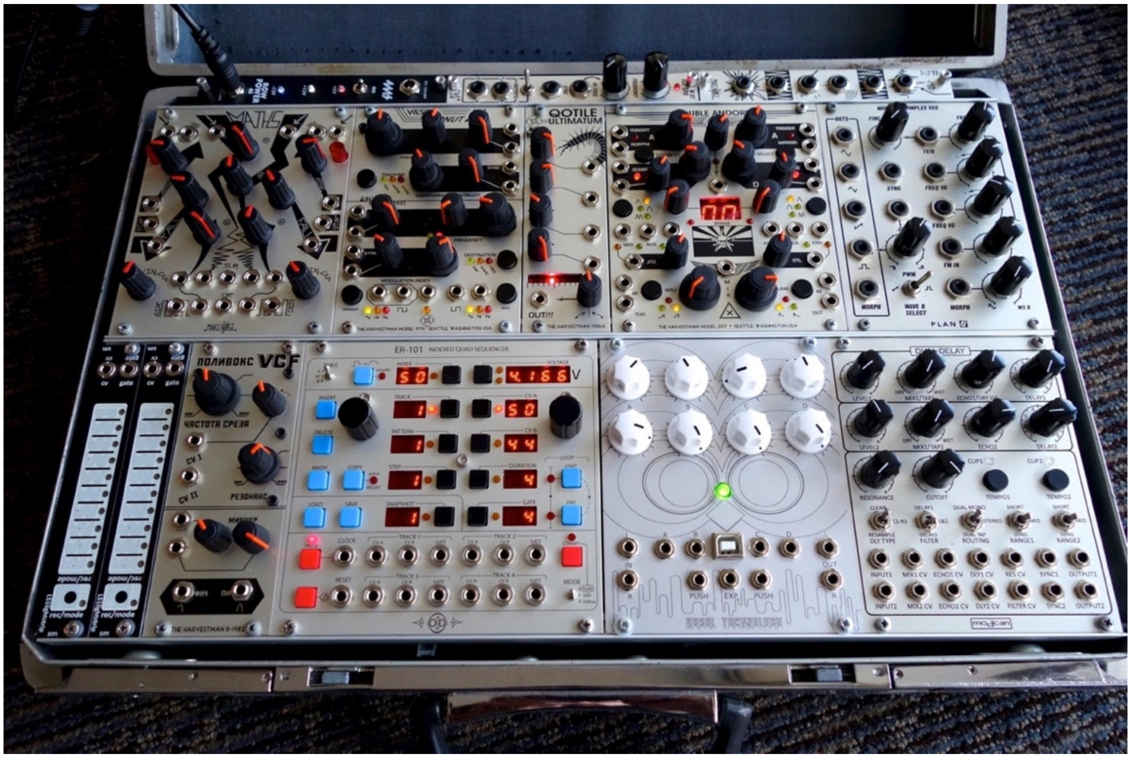


Figure 4.7. The Ear Tone Toolbox on the OWL eurorack module (bottom row, second from right). The module allows for voltage control over the various parameters of each instrument in the toolbox.

the OWL is a standalone pedal or a eurorack module for programming effects or instruments. Both versions contain a STM32F4 microcontroller with a 168 MHz 32 bit ARM Cortex M4, 192 kB RAM, 1MB of flash memory and a sampling rate adjustable up to 96 kHz. The eurorack module, as seen in Figure 4.7, allows for control voltage (CV) control over each instrument parameter in the toolbox. Given the form factor of the hardware, the multichannel version of *DPSpec* is unavailable, and the primary tone spectrum is generated with a reduced number of sine tone oscillators. Apart from *DPSpec,* the rest of the instruments in the *Ear Tone Toolbox* run on the OWL similarly to their software-only counterparts, with a few platform-specific variations. The four potentiometers on the panel allow for various methods of control depending on the instrument variation. For example, the first two knobs control the QDT and CDT in the *DiffTone* module, and the remaining two potentiometers are used in a second (dual) version of the same instrument. Other uses of the extra knobs include amplitude control for the guide tones.

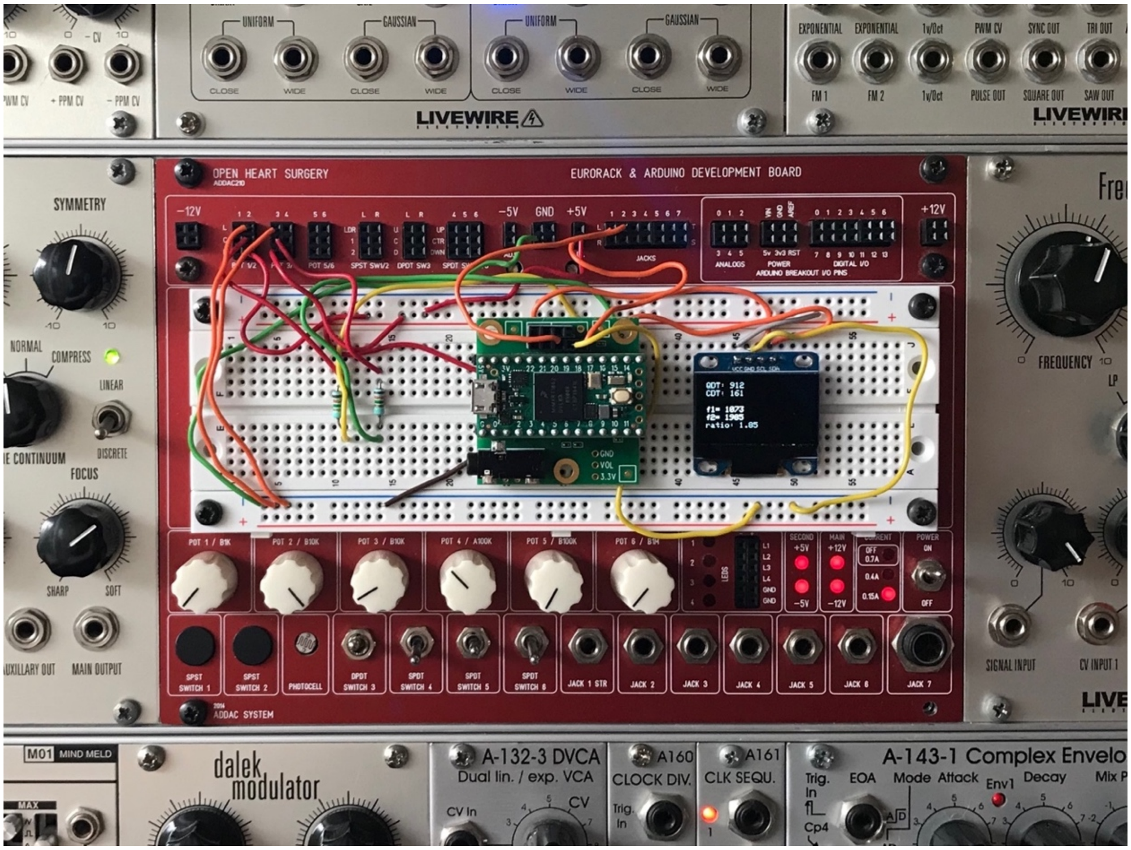


Figure 4.8. DiffTone eurorack module prototype using the Teensy 4.0, Audio Shield rev D, and 0.96" OLED display mounted on an ADDAC Systems development board.

Additionally, a build of the ETT is in development for the Teensy 4.0. The Teensy is a low-cost USB microcontroller suitable for high quality digital signal processing with a dedicated audio shield providing outputs. The Teensy 4.0 has a Cortex-M7 processor with a 600 MHz clock. With additional circuits allowing voltage control of the parameters, the Teensy version can integrate into any existing modular synthesizer format. Figure 4.8 shows a prototype running the *DiffTone* unit generator, mounted and powered in a eurorack modular synth. Two potentiometers control the respective QDT and CDT frequencies, while the audio output generates the f_1 and f_2 primaries. A benefit of the Teensy version is the OLED display, which shows the exact QDT and CDT frequencies, the respective f_1 and f_2 primaries, and their ratio.

4.4 CONCLUSION AND FUTURE DEVELOPMENT

Synthesizing distortion products with generalized unit generators enables the user to flexibly employ fundamental electronic music techniques and create larger instrument systems with a high level of creative freedom. Bouche *et al*. (2017) details how the toolbox can be used in a dynamic computer-aided composition framework. Kendall *et al.* (2014) provides an overview of select distortion product synthesis techniques, with audio examples. Future versions of the ETT will split into two paths: the continued development of low-level unit generators, and higher-level instruments with dedicated functions. In development are several unit generators built upon the results found in the perceptual work detailed in Chapter 3, particularly with three-tone stimulus. Additionally, a method for directly incorporating the most salient ratios is forthcoming. Unit generators of higher order difference tones are also underway, such as the quintic distortion product $3f_1$-$2f_2$. New instruments will be built using special purpose and advanced synthesis techniques. Such instruments will be built on completed and future perceptual research (e.g., more than three primaries), and will directly connect to/emerge from compositional techniques and needs. Finally, the hardware modular synthesis version of the Ear Tone Toolbox will continue development with a potential manufactured module release.

The Ear Tone Toolbox is the first widely available software package for auditory distortion product synthesis. The five main unit generators described in this chapter comprise the earliest versions of the toolbox, and regular updates of additional generators are planned. By releasing the software as open-source, the author intends to encourage the future development of the field of distortion product synthesis, and to provide educational tools for listening and understanding the fundamentals of combination tones through an accessible package. The software is available on the CCRMA Ear Tone Toolbox website.[4]

[4] https://ccrma.stanford.edu/~chechile/eartonetoolbox/main.html

CHAPTER 5

Practical Applications of Difference Tones in Electronic Music Composition and Synthesis

5.1 INTRODUCTION

Named after Helmholtz's seminal text, *On the Sensations of Tone* is a series of compositions that explores the physicality of sound and spatial depth through the use of auditory distortion products. With ongoing development since 2010, the series consists of nine main entries and one prelude. A parallel series, *The Ear Tone Etudes* applies findings from the study detailed in Chapter 3 to music composition. In this chapter, the pieces in both series are used as case studies for the practical application of difference tones in electronic and electroacoustic music. Before addressing each piece, a general approach to and benefits of composing with the phenomenon is provided.

5.2 UNIQUE PROPERTIES OF DIFFERENCE TONES & BEST PRACTICES

Because an acoustic layer of primaries in a space generates an internal spectrum of sound within the head, the phenomenon of difference tones is inherently spatial. It is clear the acoustic tones originate in the room by speakers or instruments, but the difference tones are localized within the head. Furthering this spatial dimension is a distinction in physical sensations between the external primary tones and the internal ear tones. The acoustic tones are not unlike most sounds one encounters, although often they occur in a sensitive frequency region. However, the internal distortion products possess a unique characteristic that Helmholtz (1954) described as a "mechanical tingling in the ear." The combination between the two sensations, as well as the two locations of the sound, contribute to an immersive experience that is unique to distortion products.

Dispersing each of the primary tones among a multichannel speaker array will increase the spatiality further. When multiple groups of primary tones are layered within the sound system, the result produces a dense spectrum of external and internal sound. Physical movement will cause ear tones to appear, disappear, and change timbre when different combinations of primaries reach the ear. The resulting experience is reactive because the listener can change the piece through head movement, or by moving to a different physical location within the space. Small changes in position can lead to entirely new perceived spectral components.

Additionally, difference tones allow the composer, musician, or sound artist access to increased harmonic content. By using primary tones known to reliably evoke particular distortion products, difference tones extend available pitch content from acoustic tones to a combination that includes ear tones. The relationship between primaries and distortion products can be programmed and automated to allow for augmentation of tonality or to support advanced synthesis techniques.

All of the properties described in this section contribute to a unique mode of listening. The dynamic between the acoustic frequencies and the resulting difference tones creates a sonic environment where listeners can shift attention between a microscopic layer of either ear tones or primaries, and a macroscopic layer of the composite sonic environment. The intensity of the difference tones can be amplified by placing slightly cupped hands near the ears, effectively extending the outer ear. Experimenting with hand placement around the ear (along with head position in the acoustic space) will yield different results and further contribute to the reactive experience of listening to difference tones (see Figure 5.1).

As the experience of auditory distortion products is unfamiliar to most, a few best practices can help create optimal listening by setting expectations. Providing liner notes, or a pre-concert talk with select details about difference tones and what to expect will pique general interest. Also explaining that cupping hands over the ears will intensify the experience for the adventurous, and the option to plug one's ears is always available if the listener is uncomfortable with the effects. Regarding technical best practices, an added benefit of splitting the primaries between speakers is the reduced chance that nonlinearities will be generated in the external sound system. Using pre-recorded sections that evoke difference tones can be an effective strategy for it allows the performer to adapt the material to the acoustics of the venue. Focusing on the parameters of presentation can prevent uncomfortable amplitude levels and create optimal conditions for evoking difference tones. Additionally, to evoke the most effective distortion products, ensure the frequency spectrum is free of components that would otherwise mask the difference tones.

Figure 5.1. A listener experimenting with hand placement to enhance the intensity of difference tones in a multichannel sound environment.

5.3 ON THE SENSATIONS OF TONE

This section outlines the entries of *On the Sensations of Tone* finished to date. The series consists of nine main entries and one prelude. Each piece in the series relies on the listener to create a discrete layer of the sound that is not acoustically present in the space. Most compositions in the series are multichannel, with the acoustic primary tones dispersed among the speaker array. The resulting acoustic spectrum evokes different combination tones depending on the physical location of the listener within the environment. Several of the compositions were presented in conferences as papers or concert pieces over the past five years, hence this section is built upon excerpts of the corresponding presentations (Chechile, 2013, 2015, 2016a, 2016b).

The first four pieces and the prelude were written between 2010 and 2013, prior to the start of my doctoral studies. The prelude to the series was performed live on October 10th, 2010 at the Orpheum Theater in Boston, Massachusetts as a supporting act for the band Primus. With a eurorack modular synthesizer, a neurobiofeedback system (to be described in the next paragraph), and a large line array stereo loudspeaker system, the piece evoked difference tones using measured voltages stored in an analog sequencer. A delay effect was later added to blur the primary tone frequency pairings, which resulted in a spectrum of ear tones generated from more than two primaries. The piece is approximately ten minutes in duration, of which two minutes contain difference tones.

The first piece of the main series was commissioned in 2010 by Issue Project Room in Brooklyn, New York and debuted the following year. The composition was performed on a 19.2 channel sound system with 15 semi-hemispherical speakers hung from the ceiling above the audience. The instrumentation consisted of a 1978 Serge Modular Music System, a eurorack modular synthesizer, a computer running Max/MSP, and an electroencephalogram (EEG). The neurobiofeedback system collected EEG readings in real-time and used the magnitude of energy within particular brainwave frequency bands to provide control voltages for the modular synthesizers (Chechile, 2007). While the structure of the piece was fixed, the neurobiofeedback introduced an improvisatory element that influenced how the work unfolded. Two seven-minute sections evoked distortion products using a variety of methods. Voltage-controlled oscillators tuned by a multimeter offered near-precise calculation of difference tones, however this method was cumbersome to set up. Often in parallel with tuned frequencies, another method used oscillators tuned within a close range, but the exact frequencies fluctuated rhythmically. The combination of techniques created a dense spectrum of both planned and unexpected ear tones. The individual primaries were dispersed between the channels of the overhead sound system. The first entry is the longest in the series at 45-minutes in duration.

On the Sensations of Tone II debuted in December of 2011 at Tufts University in Somerville, Massachusetts. The instrumentation and structure were similar to the first entry, however the piece used four channels, and consisted of three approximately five-minute sections that totaled an overall duration of 14 minutes. The middle section of the piece evoked difference tones using borrowed material from the first entry. The length and alternating three-part structure carried over to most of the other pieces in the series because the duration of ear tone material was found to be comfortable, and the contrasting material frames the difference tone sections.

On the Sensations of Tone III was performed in 2012 at Vaudeville Park in Brooklyn, New York, and was the first (of two) stereo entries in the main series. Like the previous two pieces, the distortion products were evoked by measuring control voltages on a modular synthesizer. This piece was also 14 minutes in duration, and it shared the alternating three-part structure with the previous entry. While the difference tones were audible in the small venue, the lack of a multichannel sound system created a less immersive experience. This was the last entry in the series to evoke distortion products with a modular synthesizer alone.

On the Sensations of Tone IV debuted at the 2013 Deep Listening Arts/Science Conference at Rensselaer Polytechnic Institute in Troy, NY, and was built using a prototype of the software that became the *Ear Tone Toolbox* (Chechile, 2013). The 20-minute piece was performed live on a modular synthesizer, however the middle section evoked difference tones using two custom sequencers written in Max/MSP. Influenced by the neurobiofeedback system, the timing of the two sequencers gradually shifted to evoke phasing distortion products. Similar to the *DiffTone* object, the values programmed into the sequencer corresponded to the QDT, and the software generated the necessary acoustic primary tones. This piece was presented with a four-channel sound system, and featured an additional four channels of highly-directional ultrasonic speakers that were positioned strategically in the space. When the listener was in a direct path of an ultrasonic speaker, another layer of difference tones would emerge.

The first entry written at Stanford, *On the Sensations of Tone V* debuted in December of 2013 at the Bing Concert Hall. The eight channels of audio were doubled to match 16 speakers in the venue. The first section was performed live on a modular synthesizer using the neurobiofeedback system, and did not involve difference tones. The second section evoked distortion products using custom software written in the ChucK programming language. Clusters of short repeating and sustained primary tones were layered to evoke a spectrum of difference tones that evolved as the repeating patterns phased. The second section is eleven minutes of the sixteen-minute total duration.

Also written in the ChucK programming language, *On the Sensations of Tone VI* was performed at CCRMA in March, 2014. The piece provided the audience agency to craft their own experience by physically interacting with the sound. Using eight channels, the first five minutes were built from eight pairs of primary tones that gradually change. The pace was deliberately slow, and movement was minimal to encourage the audience to actively construct their own experience in the concert space. By changing position, experimenting with hand movement around the ears, or through slight head movement, different combinations of primary tones interacted within the listener's ears and created different combinations of distortion products. The piece concluded with two minutes of contrasting fast-paced material. The instructions to explore the work through movement was expressed in both the program notes as well as a brief introduction. The composition is seven minutes in duration, and it is the only entry in the series that consists entirely of difference tones.

On the Sensations of Tone VII was written and performed at CCRMA during the spring of 2014. Like the previous piece, the audience was encouraged to engage with the material through movement or by changing head position. The eight-channel composition is 15 minutes in duration. The piece was detailed in a paper presented at the 2015 International Conference for Auditory Display (ICAD) in Graz, Austria, and performed at the same conference. The following text is an analysis of the work adapted from the publication (Chechile, 2015).

The sections of *On the Sensations of Tone VII* alternate between non-difference tone material generated with a modular synthesizer and difference tone sections generated by custom software built in Max/MSP and ChucK. The first section does not involve ear tones and serves to explore the spatiality of a traditional eight-channel sound system. The second section introduces ear tones to further expand the spatial dimension. The third section consists of a low frequency drone that acts as contrasting material in pitch, rhythm, and spatiality for the stereo signal sent to all channels. The fourth section once again expands the spatial dimension through the use of primary tone clusters that evoke a moving spectrum of distortion products.

On the Sensations of Tone VII was built from a custom non-repeating scale system derived from particularly salient quadratic difference tones that were selected from sweeping the frequency spectrum while holding a constant cubic difference tone. Each note of the scale produces two acoustic frequencies (f_1 and f_2) and the resulting quadratic and cubic difference tones. The balance between the dominant difference tone varies across the scale, however the QDT is often the most prominent. Based on the scale, polyphony was crafted to generate a complex spectrum of distortion products. Table 5.1 shows the combination tone scale system used in the composition.

note	f1	f2	QDT	CDT
1	1289	1378	89	1,200
2	1301	1402	101	1,200
3	1339	1478	139	1,200
4	1349	1498	149	1,200
5	1376	1552	176	1,200
6	1446	1692	246	1,200
7	1464	1728	264	1,200
8	1514	1828	314	1,200
9	1560	1920	360	1,200
10	1600	2000	400	1,200
11	1650	2100	450	1,200
12	1709	2218	509	1,200
13	1756	2312	556	1,200
14	1828	2456	628	1,200
15	1850	2500	650	1,200
16	1890	2580	690	1,200
17	1923	2646	723	1,200
18	1960	2720	760	1,200
19	2029	2858	829	1,200
20	2081	2962	881	1,200
21	2110	3020	910	1,200
22	2142	3084	942	1,200
23	2304	3408	1,104	1,200
24	2375	3550	1,175	1,200
25	2480	3760	1,280	1,200
26	2518	3836	1,318	1,200
27	2634	4068	1,434	1,200
28	2689	4178	1,489	1,200
29	2728	4256	1,528	1,200
30	2814	4428	1,614	1,200
31	2870	4540	1,670	1,200
32	2947	4694	1,747	1,200
33	3017	4834	1,817	1,200

Table 5.1. Scale system of acoustic stimulus pairs and associated quadratic and cubic difference tones in Hz. The cubic difference tone was held at a constant frequency to limit the combined spectrum, to emphasize the quadratic difference tone, and to create a common pitch linking the notes of the scale.

Figure 5.2. The second section of On the Sensations of Tone VII, showing the sustained chords and short bursts of combination tones. The image is a screenshot of the composition as recorded and arranged in Ableton Live.

The ear tone material is categorized into two groups: sustained chords, and short bursts. Each of these two categories overlap at times, resulting in additional primary tone combinations. Distortion products first appear in the second section of the piece, which begins with sustained primary tones that evoke chords of distortion products (see Figure 5.2). Each chord consists of eight acoustic tones (four notes from the scale) for eight loudspeakers. The resulting spectrum is complex as it contains the quadratic difference tone and cubic difference tone of each note, as well as additional distortion products produced through interactions with the other notes. The distortion product spectrum increases in density when polychords are introduced, such as the blue/green pairing in figure 5.2. The orange chord in the figure is intended to resemble a tonic of sorts, and is frequently repeated after material with more tension.

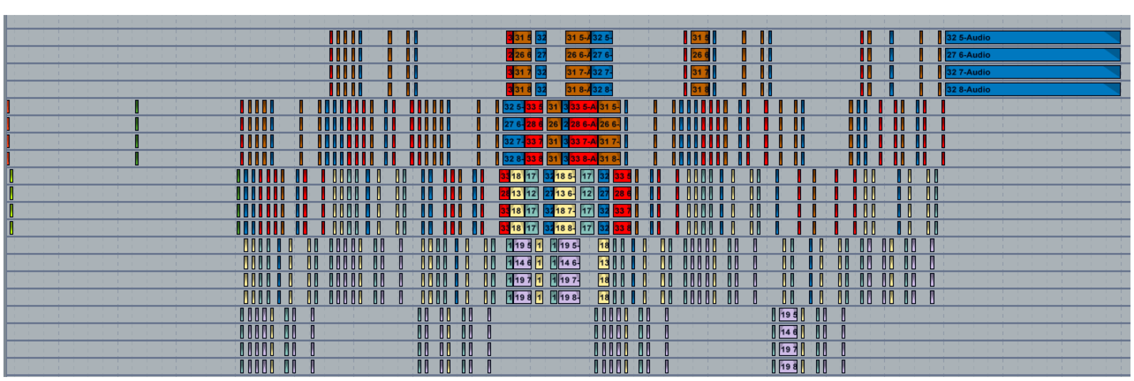


Figure 5.3. Detail for the second half of section two of On the Sensations of Tone VII.

As shown in Figure 5.3, the second half of the section combines chords from the sustained section with new chords, and introduces the sustained tones as truncated bursts. The bursts are organized into two repeating melodies that evolve through a rotating order. The two melodies play simultaneously, producing short bursts of distortion products. Duration is explored as the bursts elongate and overlap, producing different combinations of ear tone material.

The section was composed in Max/MSP by developing a keyboard that generated the ear tone scale. The distortion product chords were built through exploring ear tone harmony in an eight-channel sound system. Once the chords were selected, their duration was modified into sustained chords and short burst melodies.

The fourth section starts with the tonic chord from section two, but quickly transitions to new material. The section was composed using a custom ChucK script that generated a repeating two-chord progression of primary tones (see Table 5.2). The script ran multiple times in asynchrony to generate rhythm that phased upon each repetition. The result is a complex distortion product spectrum containing the four primary tone pairs of each chord and their respective quadratic difference tone, cubic difference tone, and internal interactions. The amplitude of each channel was adjusted to create an additional layer of change between the sounding chords (see Figure 5.4). The result produced a shifting dynamic between the primaries and the difference tones.

Ear Tone Chord 1

note	f1	f2	QDT	CDT
8	1514	1828	314	1,200
9	1560	1920	360	1,200
10	1600	2000	400	1,200
11	1650	2100	450	1,200

Ear Tone Chord 2

note	f1	f2	QDT	CDT
12	1709	2218	509	1,200
13	1756	2312	556	1,200
14	1828	2456	628	1,200
15	1850	2500	650	1,200

Table 5.2. Deconstruction of each of the two chords in the fourth section of On the Sensations of Tone VII. The note numbering is consistent with the scale system outlined in table 5.1, and the frequencies are displayed in Hz. The QDT and CDT are provided for each note, however the internal combinations between each note in the chord are not shown.

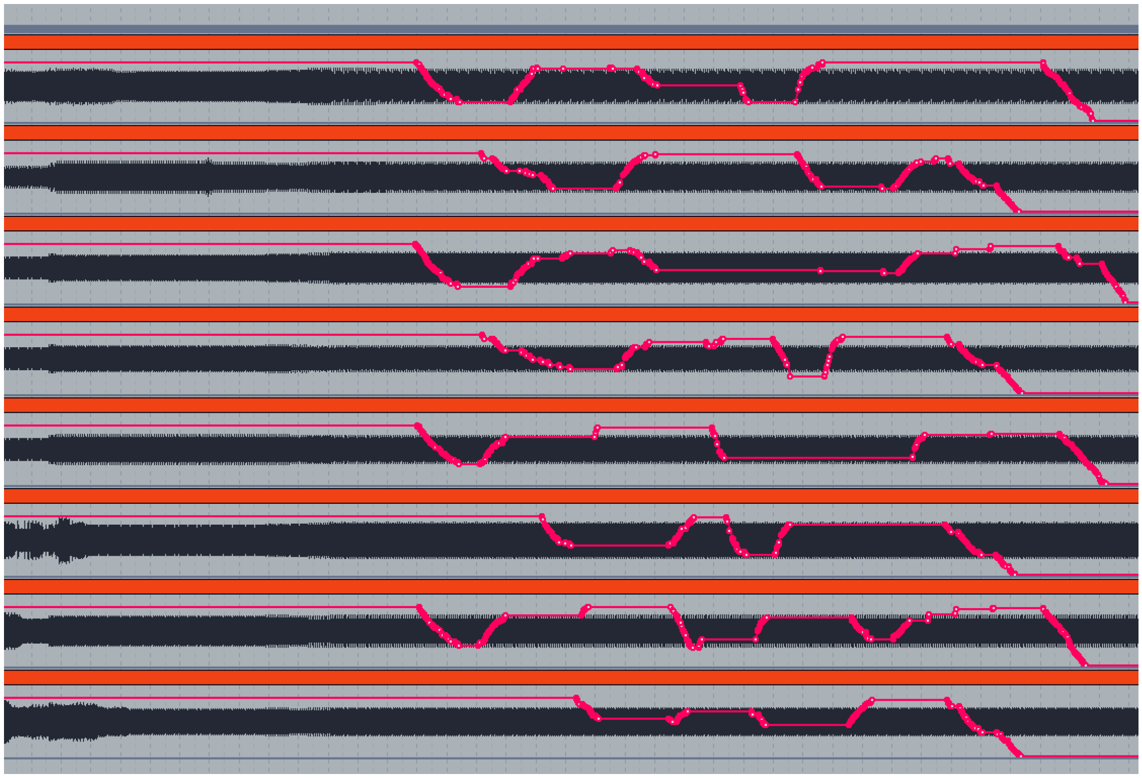


Figure 5.4. Evolving amplitude (red lines) over the composite recording of the fourth section of On the Sensations of Tone VII.

On the Sensations of Tone VIII is the first piece in the series to use acoustic instruments (Chechile, 2016b). Loren Mach and I performed the composition at the Bing concert hall in the spring of 2015. In three parts, the first section was written for the orchestral crotales with minimal (or optional) electronics, and the third section was written for electronics with optional crotales. The middle section functions as contrasting material to the difference tone sections, and was written for live and prerecorded prepared tam tams and gongs. The crotales were spatialized along with the electronics in a 24.7 channel dome with individual components of the synthesized frequencies paired to dedicated speakers.

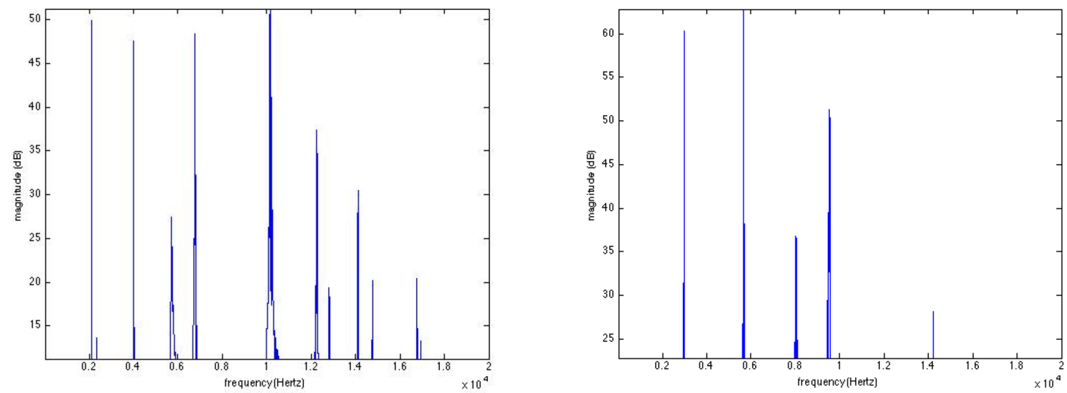


Figure 5.5. Spectral analysis conducted on the set of crotales used in On the Sensations of Tone VIII, with frequency on the x-axis and amplitude on the y-axis. The C7 note (left) shows more modes excited than the F#7 (right). The recordings were made from strikes in same relative position (3/4ths the width from the center).

In a pre-compositional process, four recordings per upper octave crotale note were analyzed for spectral content. Figure 5.5 shows the frequency components from the C7 crotale (left), which was the lowest note used, and the F#7 (right), which shows fewer modes excited. Certain crotale notes were found to contain components that produced a CDT from a single note strike (see Table 5.3). While the frequencies are generally higher than those tested in the psychoacoustic study described in Chapter 3, the ratios are within the range of the findings at lower frequencies.

Note	f1 (f20)	f2 (f30)	ratio	CDT	CDT Note
C7	2104	4008	1.90	200	G3
C#7	2226	4437	1.99	15	
D7	2359	4769	2.02	-51	
D#7	2524	5051	2.00	-3	
E7	2632	4770	1.81	494	B4
F7	2787	5010	1.80	564	C#5
F#7	2971	5659	1.90	283	C#4
G7	3132	5874	1.88	390	G4
G#7	3343	5874	1.76	812	G#5
A7	3539	6792	1.92	286	D4
A#7	3754	7278	1.94	230	A#3
B7	4025	8071	2.01	-21	
C8	4406	7848	1.78	964	B5

Table 5.3 Crotale modes of vibration, ratio, and resulting CDT with note pairing. Frequencies are in Hz, and entries in red evoke a strong CDT.

Figure 5.6 shows a scale created from the f(2,0) and f(3,0) modes of vibration, with the strongest cubic difference tones reported on the lower staff. These notes and their corresponding CDTs were deliberately used while composing. From a single crotale, the QDT values were typically either inaudible or too close to the fundamental pitch because of the component ratios. Therefore, the only QDTs explicitly considered during the compositional process were those evoked when two crotales overlapped in decay or when simultaneously struck. Appendix B provides the crotales part of the score for the first section of the piece, and Appendix C provides the first page of the crotales score re-written to explicitly include the perceived but acoustically absent CDT and QDT notes. Figure 5.7 shows a spectral analysis of the first section of the piece. The box in the lower portion of the figure surrounds the frequency region where material was generated in the listener's ears, but was absent in the acoustic space.

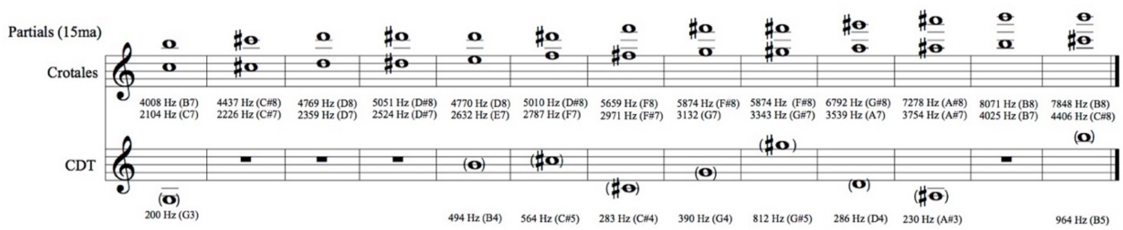


Figure 5.6. A scale created from the f(2,0) and f(3,0) modes, with the strongest cubic difference tones reported on the lower staff. The specific frequencies of the modes and the CDT are labeled.

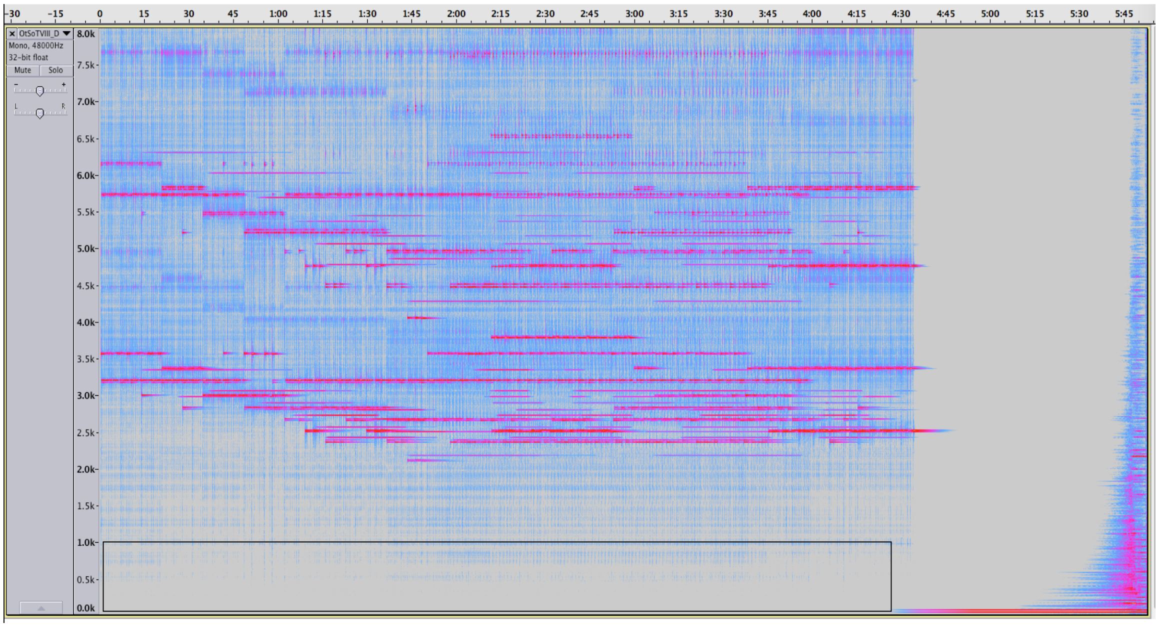


Figure 5.7. A spectral analysis of the first section of On the Sensations of Tone VIII. The lower box surrounds the frequency range of the evoked distortion products, but the frequencies are clearly absent in the acoustic space and sound recording.

On the Sensations of Tone IX: The Descent was built from field recordings made in the Paris catacombs. Standing apart from previous entries in the series, the difference tone sections were integrated into the recordings. Careful attention was made to ensure the frequency region where difference tones were evoked was free from acoustic interference of masking material. The two sections with distortion products were developed in Max/MSP using the multichannel *DPSpec* unit generator from the Ear Tone Toolbox (Chechile, 2016a). By integrating the unit generator into a larger Max patch for further processing, the resulting distortion products took on a unique character. For example, in the second instance with difference tones, the twelve individual sine waves were first amplitude modulated in unison and the resulting individual signals were modulated again in asynchrony. Slight frequency randomization was applied to both the primary tones and the specified QDT, which resulted in an uneven jitter between the acoustic tones and the evoked distortion products. The described Max patch is shown in Figure 5.8. The piece was first performed in Paris at the Cité Internationale des Arts in the Fall of 2015. Apart from the prelude, it is the second of two pieces in the series that was written specifically for stereo presentation. The piece was composed while meeting regularly with Éliane Radigue, and the content reflects the nature of our conversations at the time.

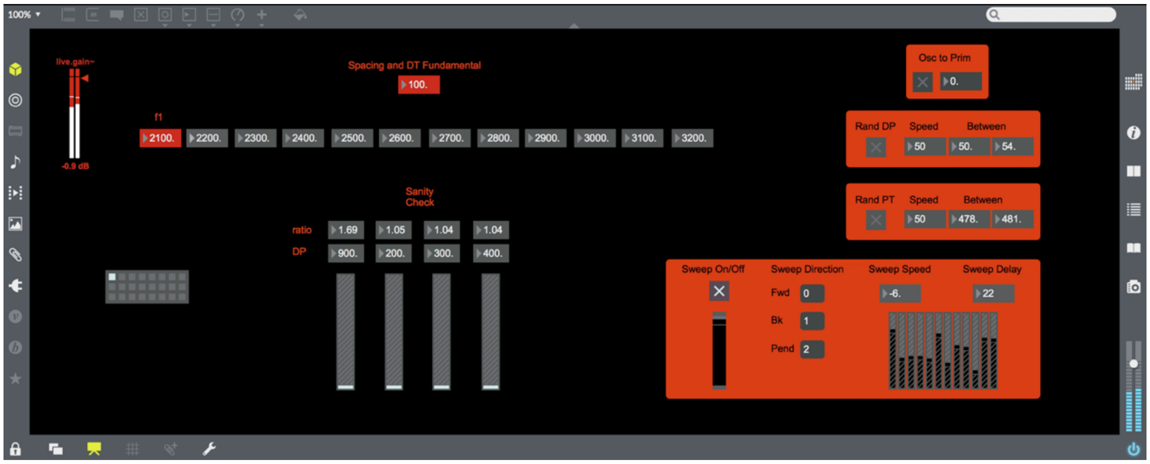


Figure 5.8. Implementation of the DPSpec Ear Tone Toolbox unit generator in a larger Max/MSP patch, as used in On the Sensations of Tone IX.

5.4 THE EAR TONE ETUDES

The Ear Tone Etudes is a collection of pieces that illustrate key findings from the psychoacoustic study described in Chapter 3. Each etude was written, programmed, and recorded in a single day. Additional time was spent on mixing, mastering, and minor edits only. The compositional process entailed focusing on a single detail from the study results and developing the piece and software around the idea. Nearly all of the audio consists of the stimulus frequencies used in the study. As will be described, after expecting a particular effect based on the data alone, the end result was often surprising. The pieces were performed at CCRMA as *Ear Tone Etudes and Interludes* on February 20th, 2020. The event was the last in-person concert at CCRMA before the COVID-19 pandemic forced widespread lockdown.

In *Ear Tone Etude 1*, the intention was to create a melody using difference tones with high discrimination indices in the $f_1+f_2-f_3$ CDT class. This particular difference tone class was chosen because it was the only class in the study that resulted from a combination of all three primary tones. As it is the first etude in the series, it functions to train the listener to differentiate between the acoustic primary tones and the internal ear tones. To achieve this goal, the difference tone melody is not initially evoked as difference tones, but as an acoustic melody with an applied reverb effect to signify distinction from the primary tones. After the repeating melody becomes familiar, the audio crossfades with primary tones that evoke the same melody, but now as true difference tones. After working out the first version of the piece, it was clear that despite having high discrimination indices, the difference tones in the melody were not entirely audible in the $f_1+f_2-f_3$ CDT class. Because the piece was to serve as an educational point of departure, and the $f_1+f_2-f_3$ tones were not consistently clear, the piece was shifted to use the difference tone class with the highest mean discrimination index and the highest percent of detection: the two-tone f_2-f_1. In the final version, the acoustic melody plays as the QDT, the CDT, and both simultaneously. The primary tones remain the same on every repeat. The total duration of the first etude is 2:20 minutes.

Ear Tone Etude 2 illustrates how a single melody can evolve as ear tone pitch content when the primary tones change in intensity. The piece begins with all three primaries at equal amplitude, and as the melody repeats, the f_3 primary decreases in amplitude, resulting in a change in difference tones. The f_3 returns to equal amplitude and the f_2 decreases, therefore producing difference tones between the f_1 and f_3 primaries. Next, the f_3 amplitude decreases until only the f_1 primary is playing, and no ear tones are generated. The changing amplitude of the primaries results in a complex shifting timbre with a spatial aspect. The length of the piece is 1:30 minutes.

Because the first two etudes were mechanical and fast paced, *Ear Tone Etude 3* was designed to have space with a relaxed, ebbing and flowing time structure while using the three-tone primaries that yield the largest number of difference tones. A secondary interest was to employ a reverberation with a long decay time to facilitate combinations between blocks of stimulus ‘chords.’ The space between each chord was gradually reduced, which resulted in multiple primary tone clusters to overlap and create additional complexity. While not entirely unexpected, the individual difference tones were not all audible, but rather, they fused to create a complex ear tone timbre. This finding was not clear when examining the data alone, but it became obvious in a musical context. The third piece in the series is three minutes in duration, and initially premiered ahead of *Ear Tone Etudes and Interludes* at a CCRMA concert in the Bing Concert Hall on February 1st, 2020.

As an aside, the multichannel convolution reverb employed in *Ear Tone Etude 3* was built using the acoustics of ancient caverns in Longyou, China. In 2017, a team of researchers from CCRMA traveled to the caves to record acoustical measurements. While the decay time was long, the reverberation in *Ear Tone Etude 3* was further expanded beyond the size of the measured space to allow for greater interaction between the decaying tones and multiple difference tones.

Picking up where the third entry ended, the intention of *Ear Tone Etude 4* was to continue to explore the density of the ear tone spectrum by increasing layers of primary tone blocks. Connected by a field recording that plays throughout, bursts of primaries fill particular silences between distant sporadic percussive sounds. On specific counts, an additional three-tone complex is layered upon the previous complex until four groups of three-tone primaries are sounded at once. Each primary tone block is localized to a dedicated pair of speakers in an eight-channel configuration, from front to rear. In a brief moment shortly before the piece concludes, a frequency complex consisting of all evoked distortion products sounds in the acoustic space, simulating the frequency and timbre of the auditory distortion products. The duration of the etude is 2:08 minutes.

Discussed in Chapter 3, a non-simultaneous alternative for testing difference tones is the pulsation-threshold method. In this method, the primary tones quickly alternate with silence, and a probe tone at the frequency of the distortion product alternates in the gaps. The probe tone amplitude is controlled by the subject, and when it is at the same level as the distortion product, the DP sounds as a continuous tone. *Ear Tone Etude 5* was built from the creative application of this technique, and it explores the distortion product spectrum generated by three sets of simultaneous primaries. The first tone complex is 519 Hz and 698 Hz, the second is 803 Hz and 921 Hz, and the third is 1358 Hz, 2212 Hz, and 2489 Hz, with probe tones at the frequencies of f_2-f_1 for the two-tone complex, and f_3-f_2 for the three-tone complex.

The piece begins with all three primary tone blocks pulsing at the rate of 100ms. During the 100ms of silence, the probe tones gradually increase in amplitude until they are the same level as the primaries. When the probe tones and primaries are the same amplitude, the pulse is transformed into a continuous tone. Gradually the probe tones decrease in amplitude and the pulse returns. This opening gesture signals the approach to the piece through exaggeration, as the probe tone levels never reach the same level as the primaries in the remainder of the etude. Instead, the three probe tones continuously fluctuate in amplitude as the primaries pulse, and the resulting ear tones appear and disappear into steady tones when they are the same level as the probe tones. Because there is wide variability in individual distortion product levels, the probe tone amplitude adjustments occur slowly to account for a range of thresholds. While not always at the same point, the audience should hear pulsing distortion products turn to steady tones, and gradually reverse back to pulsing tones. The duration of the etude is seven minutes.

Ear Tone Etude 6 uses broad and narrowband noise to exhibit masking of difference tones. The piece begins with noise alone, and after approximately a minute, the primary tones 698 Hz, 803 Hz, and 921 Hz are added to the signal. Both the noise and primaries are simultaneously audible, but the difference tones are absent due to the masking (see Figure 5.9, top). Once the noise is removed and the primaries sound alone, distortion products are immediately produced (Figure 5.9, second from top). The effect is striking since the overly familiar external/acoustic sound instantly develops an internal spectrum of difference tones following the sudden removal of only one component. Additionally, the primaries are not exactly precise, but rather consist of a cluster of acoustically beating primary tones. For example, the 698 Hz primary is joined by 699 Hz, 701 Hz, and 703 Hz, with the same initial frequency spacing for the other two primaries. The resulting effect gives both the primaries and the difference tones the quality of uneven spatial movement. Next, for the first time in the series, the primary tone frequencies differ from those used in the study when they simultaneously glissando upwards at independent rates. The resulting difference tones glide both upwards and downwards depending on the difference tone class. Noise is reintroduced and removed, however upon its the third appearance, it is filtered to be above the primary tones (Figure 5.9, third from top). In this case, the difference tones are not masked and the noise, primaries, and difference tones are all audible. The lower cutoff frequency for the noise filter gradually decreases so the noise encompasses the primaries. However, this time the amplitude of the noise is lower, so the difference tones remain audible as they coexist in the same frequency region (Figure 5.9, bottom). The remainder of the piece explores shifting harmonies created with the freely moving and beating internal and external tones. The duration of *Ear Tone Etude 6* is approximately nine minutes, and it is the last piece performed in the suite.

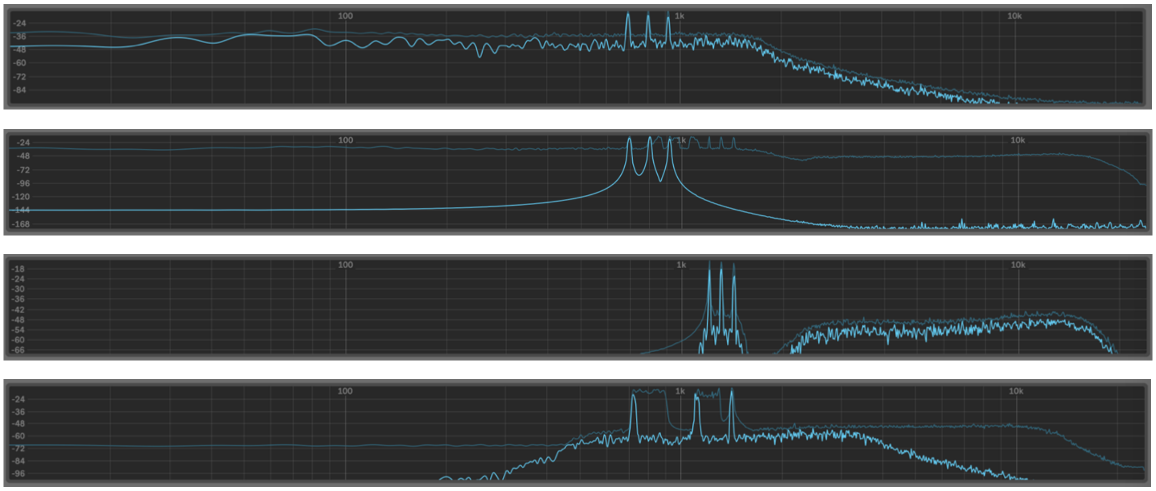


Figure 5.9. Acoustic spectrum from four moments in Ear Tone Etude 6. From top to bottom, the primary tones are seen as the three peaks just below 1kHz, however the noise band masks the resulting difference tones (top). The primaries and the ear tones are audible when the noise band is removed (second from top). The noise band occurs at a lower amplitude and frequency spectrum above the primary tones, resulting in audible distortion products (third from top). The primaries and noise band share the same frequency range, but the lower amplitude noise does not mask the difference tones (bottom).

5.5 MISCELLANEOUS WORKS

An additional composition entitled *Rides Again: For/With Pauline Oliveros* (2017) was built from combination tones of mechanical origin. In her early electronic music, Pauline Oliveros used signal generators tuned to frequencies above and below the human hearing range (Oliveros, 1967; Chechile & Thorpe, 2017). When the signals were combined and amplified, the otherwise inaudible frequencies produced perceivable tones in the equipment and speakers. Examples of work utilizing this technique include *Bye Bye Butterfly* (1965, 2012) and *I of IV* (1966, 2012). In 2006, Pauline and I began a collaboration using the aforementioned neurobiofeedback system. Although the system was primarily designed for shaping performance data in real time, we held several sessions recording our brainwave activity while listening, meditating, and playing music together. Connecting the techniques of Pauline's early electronic music to the neurobiofeedback collaboration, *Rides Again* was built exclusively from the individually inaudible EEG recordings. The previously unused brainwave files are parsed into discrete bands, scaled to frequency regions above and below the threshold of hearing, and amplified with individual gain control per band of activity. Audible tones are generated as mechanical difference tones only when our brainwaves are combined, and the piece is shaped by adjusting the amplitude of each EEG frequency band. The piece debuted at a memorial concert for Oliveros at Tufts University in April 2017. It was subsequently performed at the Bing Concert Hall at Stanford University, the Legacies of Pauline Oliveros Symposium at Brooklyn College, and the 2018 Electronic Music Midwest Festival at Lewis University.

In an untitled work-in-progress, distortion products (and/or primary tones) are pitch matched with an acoustic piano in real-time. Built using Max/MSP and the Ear Tone Toolbox, two 64-step sequencers send the pitch of difference tones to a Disklavier piano while simultaneously playing the sinusoidal primaries. With tight synchronization and controlled amplitudes, the resulting sensation resembles a timbral fusion of ear tones, piano reverberance, and mechanical noise of the Disklavier.

5.6 CONCLUSION

The compositions discussed in this section are examples of effective methods for incorporating auditory distortion products in electronic and electroacoustic music. With difference tones, the pieces possess unique properties including spectral and spatial immersion, reactive sound fields, and multilayered listening. In the following postscript is a summary of the work contained in this dissertation, and a reflection on the future directions of the work.

CHAPTER 6
Final Thoughts

6.1 POSTSCRIPT

By evoking auditory distortion products in music, the ear becomes an instrument and the act of listening reciprocates a means for performance. The spatial depth created between sounds in the room and those evoked in the ear allows for increased immersion and expanded listening practices. The work described herein would not be possible without an interdisciplinary approach involving psychoacoustic research, instrument design, and compositional need.

The present work provides the limits for detection of up to nine difference tone classes from two and three-tone primaries across a wide frequency spectrum. The research was put into practice in compositions and etudes, which were constructed with an open-source toolbox of instruments. The ability to control individual components in the distortion product spectrum allows for advanced synthesis techniques. More options for synthesis become available with known parameters for the detection of higher order difference tones, and DTs evoked from a spectrum of primary tones. Fortunately, the methods developed for the study described in Chapter 3 are scalable, which will allow for future studies using more than three primaries. The mathematical model and computer implementation can enable early predictions. Finally, all of the described work will be put to use in future compositions and installations, just as future compositional questions will fuel further research in psychoacoustic study and software design.

Appendices

APPENDIX A
Model Implementation in Mathematica

```
adpC[pow_, freqs__] :=
 Module[{freqlist, nfreq, type, displayFulleq, genform, orderlist, colorlist, mycolors, displayeqOnepower, dplist, ndp,
   allplot, displayFulleq2, vardisplayFulleq2, ndisplayFulleq2, fulleqlist, coeflist, coef, widthorderlist, sortedwidthorderlist,
   allplot2, lowstim, uptolowstimlist, belowstimplot, trig1, trig2, dont2, trig2Edit, trig3, trig3actual, trig4, trig5, dont5,
   dont5list, trig5Edit, trigIDs, cubictrigIDs, quadraticIDs, geneq, frelat, freqReplace, actual, finalcombolist, displayFulleqList,
   displayFulleqListnoTs, actualFinalCombo, f2ones, numOnes, onlyFs, numFs, onlyDiffs, onlyDiffsFreqs, erbt, band, allerbs,
   playfreqsgenform, indivfreqs, nplaydp, sampleplaylist, audioplayer, playdp, audioplaylist, onlyDiffsFreqsPlay,
   nOnlyDiffsFreqsPlay, stimfreqslist, nstimfreqs, amp, fullstimlist, stimsamplerate, playstims},

  (* on 6/10/18 this version was last saved as adpC_CustomTrigFunction_OnlyFsThatMatter_c21_erb_NoPopsDTonly.nb from 5/31/18*)

  (* calculates and plots the DPs from any number of stim tones *)
  freqlist = {freqs};
  nfreq = Length[freqlist];
  type = pow;
  For[i = 1, i ≤ nfreq, i++, f_i = freqlist[[i]]]; (* all freqs *)
  For[j = 1, j ≤ nfreq, j++, a_j = 1]; (* amplitude for each freq set to 1*)
  (*CONCATENATED VERSION*)
  displayFulleq = Expand[TrigReduce[∑_{type=1}^{type} (∑_{i=1}^{nfreq} a_i * Sin[2 π * f_i * t * (1 / (2 * π))])^type]]; (* *(1/(2*π)) for freq in Hz *)
  (*plot by type/color *)
  genform = (∑_{i=1}^{nfreq} a_i * Sin[2 π * f_i * t * (1 / (2 * π))]); (* *(1/(2*π)) for freq in Hz *)
  orderlist = {}; (* list of all freqs for all powers in Hz, ordered as linear first, quadratic,etc, and sent to allplot *)
  colorlist = {};
  mycolors = {Purple, Red, Blue, Yellow, Orange, Black, Brown, Green, White};
  For[g = 1, g ≤ type, g++,
   displayeqOnepower = Expand[TrigReduce[genform^g]];
   dplist = Variables[displayeqOnepower];
   ndp = Length[dplist];
      For[l = 1, l ≤ ndp, l++,
           AppendTo[orderlist, Part[dplist[[l]], 1, 1]];
           AppendTo[colorlist, mycolors[[g]]];
         ]
    ]; (* both AppendTos for l and g here to sort the bar chart according to group colors per power type *)
  allplot = BarChart[orderlist, ChartStyle → {colorlist}, LabelingFunction → (Placed[Panel[#1, FrameMargins → 0], Above] &),
    PlotTheme → "Detailed"];


  (* A second chart that plots amplitude as width, w/ amplitudes for each DP combined from all types (lin, quad,cubic, etc) *)
  displayFulleq2 = Expand[TrigReduce[∑_{type=1}^{type} (∑_{i=1}^{nfreq} a_i * Sin[2 π * f_i * t * (1 / (2 * π))])^type]]; (* *(1/(2*π)) for freq in Hz *)
  vardisplayFulleq2 = Variables[displayFulleq2];
  ndisplayFulleq2 = Length[vardisplayFulleq2];
  fulleqlist = {};
  For[c = 1, c ≤ ndisplayFulleq2, c++, AppendTo[fulleqlist, Part[vardisplayFulleq2[[c]], 1, 1]]];
  coeflist = {}; (* amplitude list, per each order *)
  coef = Coefficient[displayFulleq2, vardisplayFulleq2]; (* to extract coefficients *)
  For[b = 1, b ≤ ndisplayFulleq2, b++, AppendTo[coeflist, Abs[coef[[b]]]]];
  widthorderlist = Transpose[{coeflist, fulleqlist}];
  sortedwidthorderlist = Sort[widthorderlist, #1[[2]] < #2[[2]] &];
  allplot2 = RectangleChart[sortedwidthorderlist, ChartStyle → Purple, LabelingFunction → (Placed[Panel[#1, FrameMargins → 0], Above] &),
    PlotTheme → "Detailed"];


  (* a third chart that displays amplitude as width, but only gives freqs up to the lowest stim tone *)
  lowstim = Min[freqlist]; (* the lowest stim tone *)
  uptolowstimlist = {}; (* list of DPs below the lowest stim tone *)
  For[d = 1, Part[sortedwidthorderlist, d, 2] < lowstim, d++, AppendTo[uptolowstimlist, sortedwidthorderlist[[d]]]];
  belowstimplot = RectangleChart[uptolowstimlist, ChartStyle → Black, LabelingFunction → (Placed[Panel[#1, FrameMargins → 0], Above] &),
    PlotTheme → "Detailed"]; (* plot of only the DPs produced below the lowest stim *)


  (* F Ordering using custom trig IDs to solve the problem of Mathematica's cannonical ordering *)
  (* trig 1 ID *)
  trig1 = Table[Sin[F_a]^3 → (3 / 4) * Sin[F_a] - (1 / 4) * Sin[3 * F_a], {a, nfreq}];
  (* trig 2 ID *)
  trig2 = Flatten[Table[(Sin[F_a]^2) (Sin[F_b]) → (1 / 2) * Sin[F_b] - (1 / 4) * (HoldForm[Sin[2 * FA - FB]] - HoldForm[Sin[2 * FA + FB]]) /. {FA → F_a, FB → F_b},
     {a, nfreq}, {b, nfreq}]];
  dont2 = Table[Sin[F_h]^3 → 1/4 (-Sin[2 Fh - Fh] + Sin[2 Fh + Fh]) + Sin[F_h]/2 /. {Fh → F_h}, {h, nfreq}];
  (* collects the cases when F_1 and F_2 are same value *)
  trig2Edit = DeleteCases[trig2, Alternatives @@ dont2]; (* eliminates the cases when F_1 and F_2 are same value *)
  (* trig 3 ID *)
  trig3 =
   Flatten[
    Table[
```

```
    (Sin[F_a]) (Sin[F_b]) (Sin[F_c]) →
      (1/4) * (HoldForm[Sin[FA + FB - FC]] + HoldForm[Sin[FA + FC - FB]] + HoldForm[Sin[FB + FC - FA]] - HoldForm[Sin[FA + FB + FC]]) /.
     {FA → F_a, FB → F_b, FC → F_c}, {a, nfreq}, {b, nfreq}, {c, nfreq}]];
trig3actual = {};
For[v = 1, v ≤ Length[trig3], v++,
 If[Length[trig3[[v, 1]]] == 3, AppendTo[trig3actual, trig3[[v]]]]];
(* trig 4 ID *)
trig4 = Table[Sin[F_a]^2 → 1 - (1/2) * Cos[2 * F_a], {a, nfreq}]; (* adjusted ID w/o using Amplitude *)
(* trig 5 ID *)
trig5 = Flatten[Table[Sin[F_a] * Sin[F_b] → (1/2) * HoldForm[Cos[FB - FA]] - (1/2) * HoldForm[Cos[FB + FA]] /. {FA → F_a, FB → F_b},
   {a, nfreq}, {b, nfreq}]];
dont5 = Table[Sin[F_h]^2 → (1/2) * Cos[Fh - Fh] - (1/2) * Cos[Fh + Fh], {h, nfreq}];
dont5list = {};
For[k = 1, k ≤ Length[dont5], k++, AppendTo[dont5list, dont5 /. {Fh → F_k}]];
trig5Edit = DeleteCases[trig5, Alternatives @@ Flatten[dont5list]];
(* eliminates the cases when F_1 and F_2 are same value *)
(* trig IDs in groups *)
trigIDs = {trig1, trig2Edit, trig3actual, trig4, trig5Edit};
cubictrigIDs = {trig1, trig2Edit, trig3actual};
quadraticIDs = {trig4, trig5Edit};
(* apply the trig IDs to a general concatenated formula for F relationships *)
geneq = Expand[Sum_{type=1}^{type} (Sum_{i=1}^{nfreq} Sin[F_i])^type] /. Flatten[trigIDs]; (* removed Amplitude from this version *)
frelat = Variables[geneq];
(* replace the F relationships with the actual frequency values for calculated DPs *)
freqReplace = {};
For[z = 1, z ≤ nfreq, z++, AppendTo[freqReplace, F_z → freqlist[[z]]]];
actual = ReleaseHold[Variables[geneq] /. freqReplace];
finalcombolist = Table[{frelat[[i]], actual[[i]]}, {i, Length[frelat]}] ; (* f relationships and freqs,
but more listed than used in the polynomial*)
(* fit the finalcombolist options to the ones actually used in the main polynomial, which accommodates for cancellations *)
displayFulleqList = Variables[displayFulleq]; (* a list of all sins and cos with frequency and t value*)
displayFulleqListnoTs = displayFulleqList /. t → 1; (* same, but without the t, which was required to use Variables[],
but can't have for Select[] *)
actualFinalCombo = Select[finalcombolist, MemberQ[displayFulleqListnoTs, #[[2]]] &]; (* only the f relationships & freqs that matter *)


(* this block makes a list of only the difference tones *)
f2ones = Table[F_i → 1, {i, nfreq}];
numOnes = Table[Variables[actualFinalCombo][[i, 1, 1]], {i, Length[actualFinalCombo]}] /. f2ones;
(* the number of each combo if using 1's *)
onlyFs = Table[actualFinalCombo[[i, 1]], {i, Length[actualFinalCombo]}];
numFs = {}; (* the number of F's in a combo *)
For[w = 1, w ≤ Length[actualFinalCombo], w++, AppendTo[numFs, Count[onlyFs[[w]], F, Infinity]]];
(* if the numFs > numOnes, keep it and add it to onlyDiffs *)
onlyDiffs = {}; (* a list of only the difference tones and the associated numerical value *)
For[y = 1, y ≤ Length[actualFinalCombo], y++, If[numFs[[y]] > numOnes[[y]], AppendTo[onlyDiffs, actualFinalCombo[[y]]]]];

(* this block evaluates only the DTs to erb bandwidth *)
onlyDiffsFreqs = Table[onlyDiffs[[i, 2, 1]], {i, Length[onlyDiffs]}]; (* a list of only the diff freqs w/o f relationship *)
erbt = {}; (*erb transition bands *)
Do[AppendTo[erbt, (10^(i/21.4) - 1) * 1000 / 4.37], {i, 1.5, 40.5, 1}];
Do[band_i = {erbt[[i]], erbt[[i + 1]]}, {i, Length[erbt] - 1}]; (* makes each band up to erb 39 *)
(* Do[Print["band"_i," = ",band_i],{i,Length[erbt]-1}]; *) (* Print All erb Transitions *)
allerbs = {}; (* Call to get a list of bands and values *)
j = 1;
k = 1;
While[k ≤ 39,
 While[j ≤ Length[onlyDiffsFreqs],
  If[
   band_k[[1]] ≤ onlyDiffsFreqs[[j]] ≤ band_(k+1)[[1]],
     (* Print[onlyDiffsFreqs[[j]]," is in band"_k]; *) (* Prints All Band Results *)
     AppendTo[allerbs, {"band"_k, onlyDiffsFreqs[[j]]}]];
  ];
  j++;
 ];
 j = 1;
 k++;
];
```

```
(*
(* Plays the DPs with adjustable SR for faster processing *) (*CONCATENATED VERSION *)
playfreqsgenform=Expand[TrigReduce[Σ_{type=1}^{type} (Σ_{i=1}^{nfreq} a_i*Sin[f_i*(2π*t)])^type]];
indivfreqs=Variables[playfreqsgenform]; (* returns Sin or Cos[2π f t] for every freq *)
nplaydp=Length[indivfreqs];
sampleplaylist={};(* a list of sampling rates for individual Play objects *)
For[s=1,s≤ nplaydp, s++,AppendTo[sampleplaylist,Part[indivfreqs[[s]],1,1]*2.1]]; (*multiply each DP by 2.1 to accomodate Nyquist *)
(* divide by two in the audioplayer below to show Hz in the Label *)
audioplayer[playdp_]:=Labeled[Play[Evaluate[.2*indivfreqs[[playdp]]],{t,0,1},SampleRate→sampleplaylist[[playdp]],PlayRange→ {-1,1}],
  List[Abs[Part[indivfreqs[[playdp]],1,1]/2]]];
audioplaylist={}; (* a playlist of individual Play functions *)
For[m=1,m≤ nplaydp,m++,AppendTo[audioplaylist,audioplayer[m]]];
*)

(* Plays ONLY THE DTs with adjustable SR for faster processing *) (*CONCATENATED VERSION *)
onlyDiffsFreqsPlay = {};
Table[AppendTo[onlyDiffsFreqsPlay, Sin[onlyDiffsFreqs[[i]] * 2 * Pi * t]], {i, Length[onlyDiffsFreqs]}];
nOnlyDiffsFreqsPlay = Length[onlyDiffsFreqsPlay];
sampleplaylist = {}; (* a list of sampling rates for individual Play objects *)
For[s = 1, s ≤ nOnlyDiffsFreqsPlay, s ++, AppendTo[sampleplaylist, Part[onlyDiffsFreqsPlay[[s]], 1, 1] * 2.1]];
(*multiply each DP by 2.1 to accomodate Nyquist *)
(* divide by two in the audioplayer below to show Hz in the Label *)
fade = 0.04;
audioplayer[playdp_] :=
 Labeled[Play[Evaluate[LogisticSigmoid[(5 t) / fade - 5] LogisticSigmoid[(5 (2 - t)) / fade - 5] .2 * onlyDiffsFreqsPlay[[playdp]]],
   {t, 0, 2}, SampleRate → sampleplaylist[[playdp]], PlayRange → {-1, 1}], List[Abs[Part[onlyDiffsFreqsPlay[[playdp]], 1, 1] / 2]]];
audioplaylist = {}; (* a playlist of individual Play functions *)
For[m = 1, m ≤ nOnlyDiffsFreqsPlay, m ++, AppendTo[audioplaylist, audioplayer[m]]];


(* Generates the stimulus tones as a single Play[], with adjustable SR for faster processing *)
stimfreqslist = {freqs};
nstimfreqs = Length[stimfreqslist];
amp = (1 / nstimfreqs) - .1;
fullstimlist = {}; (* list of all stimulus freqs, populated by the For loop below *)
For[o = 1, o ≤ nstimfreqs, o ++, AppendTo[fullstimlist, amp * Sin[2 π * stimfreqslist[[o]] * t]]];
stimsamplerate = Max[stimfreqslist] * 2.1; (* multiply the max stimulus freq by 2.1 for Nyquist for individual SR *)
(* IMPORTANT: use function Evaluate below to avoid issue with multiple tracks in a single Play function *)

fadestim = 0.04;
playstims = Labeled[Play[Evaluate[LogisticSigmoid[(5 t) / fadestim - 5] LogisticSigmoid[(5 (2 - t)) / fadestim - 5] fullstimlist],
   {t, 0, 2}, SampleRate → stimsamplerate, PlayRange → {-1, 1}], stimfreqslist];

Print["Full Polynomial: "];
Print@displayFulleq;
Print[" "];
Print["Only the Difference Tones: "];
Print[onlyDiffs];
Print[" "];
Print["DTs in erb bands: "];
Print@allerbs;
Print@" ";
Print@"All F relationships: ";
Print[actualFinalCombo];
Print[" "];
Print@"All Freqs plotted by type: ";
Print@allplot;
Print[" "];
Print["All Freqs Plotted by Freq/Amp (width): "];
Print@allplot2;
Print[" "];
Print@"freqs below stim tones: ";
Print@belowstimplot;
Print[" "];
Print@"stim freqs: ";
Print@playstims;
Print@"DT freqs: ";
Print@audioplaylist;
Print["DTs in erb bands (again): "];
Print@allerbs;

]

adpC[3, 698, 921, 1051]
```

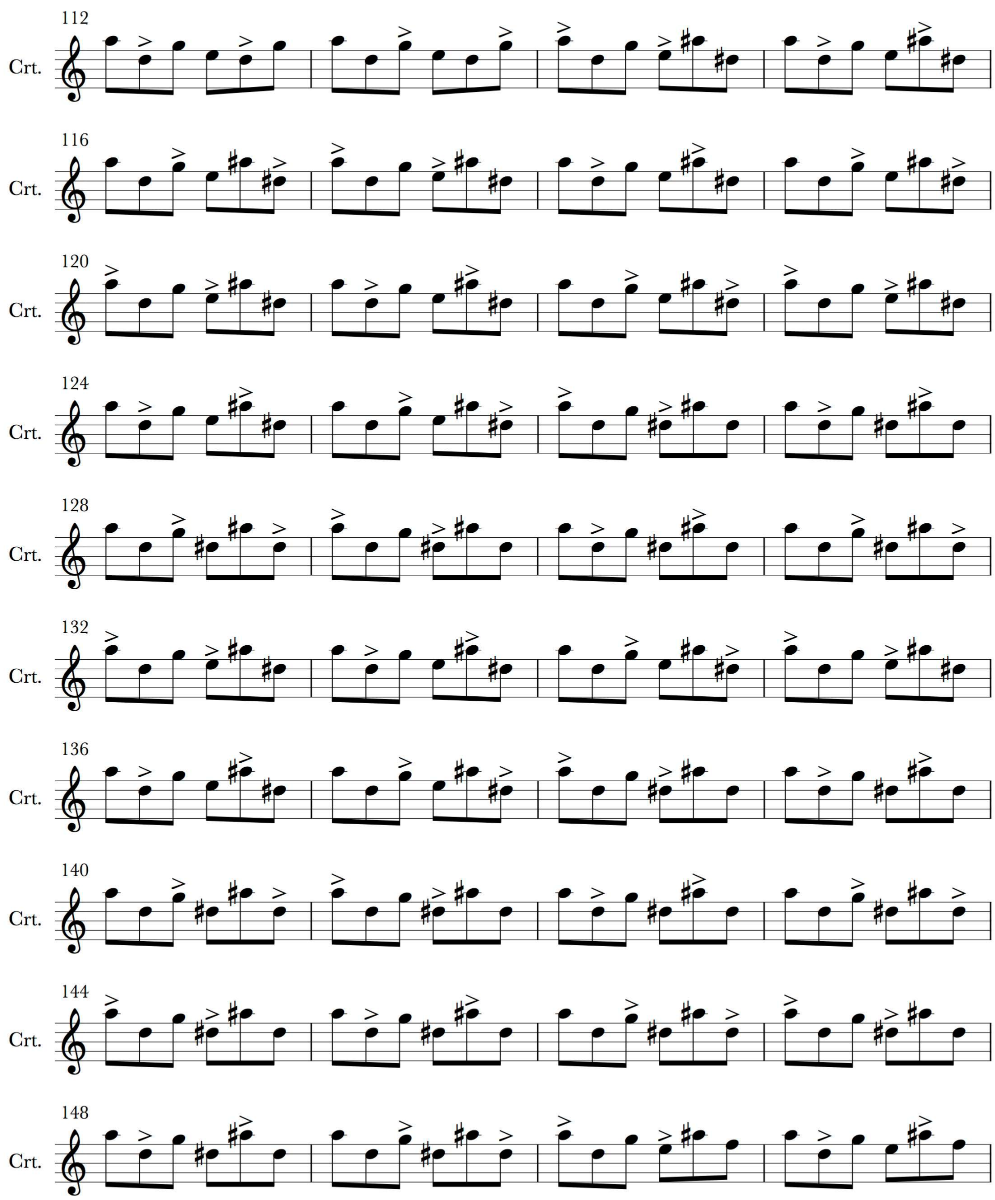
APPENDIX B
On the Sensations of Tone VIII (First Section)

On the Sensations of Tone VIII

Alex Chechile

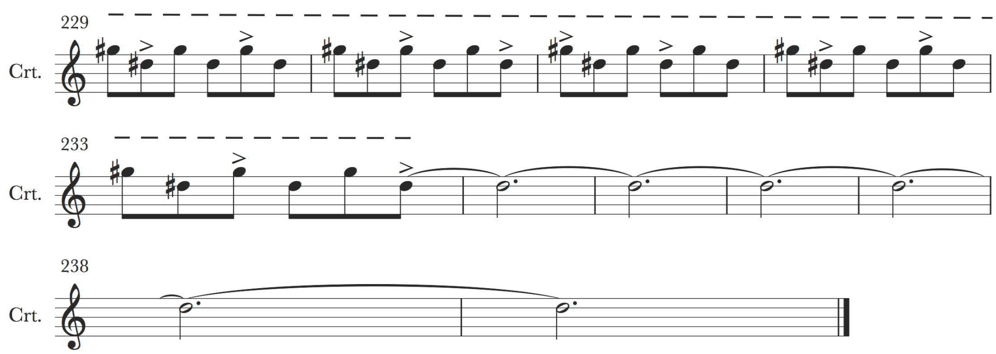

33
Crt.
37
Crt.
41
Crt.
45
Crt.
49
Crt.
53
Crt.
57
Crt.
61
Crt.
65
Crt.
69
Crt.


73
Crt.
77
Crt.
81
Crt.
85
Crt.
89
Crt.
93
Crt.
96
Crt.
100
Crt.
104
Crt.
108
Crt.


112
Crt.
116
Crt.
120
Crt.
124
Crt.
128
Crt.
132
Crt.
136
Crt.
140
Crt.
144
Crt.
148
Crt.


152
Crt.
156
Crt.
160
Crt.
164
Crt.
168
Crt.
171
Crt.
174
Crt.
178
Crt.
182
Crt.
185
Crt.


189
Crt.
193
Crt.
197
Crt.
201
Crt.
205
Crt.
209
Crt.
213
Crt.
217
Crt.
221
Crt.
225
Crt.
rit.


229
Crt.
233
Crt.
238
Crt.

APPENDIX C
On the Sensations of Tone VIII (First 40 Measures)

The following pages consist of the first 40 measures of *On the Sensations of Tone VIII* with explicitly notated cubic and quadratic difference tones, but the tones are generated in the ear and not performed directly.

On the Sensations of Tone VIII

Alex Chechile


9
Crt.
QDT
CDT


13
Crt.
QDT
CDT


17
Crt.
QDT
CDT
21
Crt.
QDT
CDT


25
Crt.
QDT
CDT
29
Crt.
QDT
CDT


33
Crt.
QDT
CDT
37
Crt.
QDT
CDT

References

Amacher, Maryanne. 1994. “Synaptic Island: A Psybertonal Topology.” In *Architecture as a Translation of Music*, edited by Elizabeth Martin. Pamphlet Architecture 16. New York: Princeton Architectural Press.

Amacher, Maryanne. 1999. *Sound Characters (Making the Third Ear)*. Compact Disc. Tzadik.

Ashmore, Jonathan. 2000. “Hearing.” In *Sound*. Cambridge University Press.

Avan, P., B. Buki, and C. Petit. 2013. “Auditory Distortions: Origins and Functions.” *Physiological Reviews* 93 (4): 1563–1619.

Baldwin, Stacey M., Byron J. Gajewski, and Judith E. Widen. 2010. “An Evaluation of the Cross-Check Principle Using Visual Reinforcement Audiometry, Otoacoustic Emissions, and Tympanometry.” *Journal of the American Academy of Audiology* 21 (03): 187–96. https://doi.org/10.3766/jaaa.21.3.7.

Bian, Lin, and Nicole M. Scherrer. 2007. “Low-Frequency Modulation of Distortion Product Otoacoustic Emissions in Humans.” *The Journal of the Acoustical Society of America* 122 (3): 1681. https://doi.org/10.1121/1.2764467.

Boer, E. de. 1956. “Pitch of Inharmonic Signals.” *Nature* 178 (4532): 535–36. https://doi.org/10.1038/178535a0.

Boring, Edwin G. 1942. *Sensations and Perception in the History of Experimental Psychology*. First Edition. U.S.A.: D. Appleton-Century Company.

Bouche, Dimitri, Jérôme Nika, Alex Chechile, and Jean Bresson. 2017. "Computer-Aided Composition of Musical Processes." *Journal of New Music Research* 46 (1): 3–14. https://doi.org/10.1080/09298215.2016.1230136.

Braunmühl, H.J. von, and W. Weber. 1937. "Ueber Die Störfähigkeit Nichtlinearer Verzerrungen." *Akust. Z.*, no. 2: 135–47.

Brown, Ann M., and David T. Kemp. 1984. "Suppressibility of the $2f_1$-f_2 Stimulated Acoustic Emissions in Gerbil and Man." *Hearing Research* 13 (1): 29–37. https://doi.org/10.1016/0378-5955(84)90092-3.

Brownell, W. E. 1990. "Outer Hair Cell Electromotility and Otoacoustic Emissions." *Ear and Hearing* 11 (2): 82–92.

Buunen, T. J. F. 1975. "Two Hypotheses on Monaural Phase Effects." *Acustica* 34 (2): 98–105.

Buunen, T. J. F., J. M. Festen, F. A. Bilsen, and G. van den Brink. 1974. "Phase Effects in a Three-component Signal." *The Journal of the Acoustical Society of America* 55 (2): 297–303. https://doi.org/10.1121/1.1914501.

Buunen, T.J.F., and F.A. Bilsen. 1974. "Subjective Phase Effects and Combination Tones." In *Facts and Models in Hearing*, 8:344–52. Communication and Cybernetics. Tutzing, Oberbayern, Germany: Springer-Verlag.

Campbell, Murray, and Clive A. Greated. 1988. *The Musicians' Guide to Acoustics*. 1st American ed. New York: Schirmer Books.

Chechile, Alex. 2007. "Music Reinformed by the Brain." Troy, NY: Rensselaer Polytechnic Institute.

Chechile, Alex. 2013. "Composing with Otoacoustic Emissions, Ultrasonic Speakers, and Neurobiofeedback." In *Deep Listening: Art/Science Conference*. EMPAC, Troy, NY.

Chechile, Alex. 2015. "Creating Spatial Depth Using Distortion Product Otoacoustic Emissions in Music Composition." In *ICAD in Space: Interactive Spatial Sonification*, 50–53. Graz, Austria.

Chechile, Alex. 2016a. "The Ear Tone Toolbox for Auditory Distortion Product Synthesis." In *Proceedings for the International Computer Music Conference*, 519–23. Utrecht, The Netherlands.

Chechile, Alex. 2016b. "The Perception of Auditory Distortion Products from Orchestral Crotales." In *Proceedings of the 14th International Conference on Music Perception and Cognition*. San Francisco, CA.

Chechile, Alex. 2017. "Evaluation of Complex Stimuli and Resulting Distortion Product Spectrum in Auditory Distortion Product Synthesis." In *The Journal of the Acoustical Society of America*. Vol. 142, No. 4. Pt. 2 of 2. New Orleans, LA: Acoustical Society of America through AIP Publishing LLC.

Chechile, Alex, and Suzanne Thorpe. 2017. "Live Performance Considerations for Pauline Oliveros' Early Electronic Music." In *Still Listening: A Series of Events in Memory of Pauline Oliveros*. McGill University.

Chertoff, Mark E., and Kurt E. Hecox. 1990. "Auditory Nonlinearities Measured with Auditory-evoked Potentials." *The Journal of the Acoustical Society of America* 87 (3): 1248–54. https://doi.org/10.1121/1.398800.

Cianfrone, G., G. Altissimi, M. Cervellini, A. Musacchio, and R. Turchetta. 1994. "Suppression Tuning Characteristics of $2f_1$-f_2 Distortion Product Otoacoustic Emissions." *British Journal of Audiology* 28 (4–5): 205–12. https://doi.org/10.3109/03005369409086569.

Davis, H. 1983. "An Active Process in Cochlear Mechanics." *Hearing Research* 9 (1): 79–90.

Dhar, Sumitrajit, and James W. Hall III. 2012. *Otoacoustic Emissions: Principles, Procedures, and Protocols*. Core Clinical Concepts in Audiology. San Diego, California: Plural Publishing Inc.

Fahey, P.F., B.B. Stagner, B.L. Lonsbury-Martin, and G.K. Martin. 2000. "Nonlinear Interactions That Could Explain Distortion Product Interference Response Areas." *J Acoust Soc Am* 108 (4): 1786–1802.

Fastl, Hugo, and Eberhard Zwicker. 2007. *Psychoacoustics*. Berlin, Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-540-68888-4.

Feeney, M. Patrick. 1997. "Dichotic Beats of Mistuned Consonances." *The Journal of the Acoustical Society of America* 102 (4): 2333–42. https://doi.org/10.1121/1.419602.

Fletcher, Harvey. 1929. *Speech and Hearing*. New York: D. Van Nostrand Company, Inc.

Fogiel, M. 1980. *Handbook of Mathematical Formulas, Tables, Functions, Graphs, Transforms for Mathematicians, Scientists, Engineers*. New York, NY: Research & Education Association.

Furst, M., W. M. Rabinowitz, and P. M. Zurek. 1988. “Ear Canal Acoustic Distortion at $2f_1-f_2$ from Human Ears: Relation to Other Emissions and Perceived Combination Tones.” *The Journal of the Acoustical Society of America* 84 (1): 215–21. https://doi.org/10.1121/1.396968.

Gaskill, Sally A., and Ann M. Brown. 1990. “The Behavior of the Acoustic Distortion Product, $2f_1-f_2$, from the Human Ear and Its Relation to Auditory Sensitivity.” *The Journal of the Acoustical Society of America* 88 (2): 821–39. https://doi.org/10.1121/1.399732.

Glasberg, B. R., and B. C.J. Moore. 1990. “Derivation of Auditory Filter Shapes from Notched-Noise Data.” *Hearing Research* 47: 103–38.

Gockel, Hedwig E., Redwan Farooq, Louwai Muhammed, Christopher J. Plack, and Robert P. Carlyon. 2012. “Differences between Psychoacoustic and Frequency Following Response Measures of Distortion Tone Level and Masking.” *The Journal of the Acoustical Society of America* 132 (4): 2524–35. https://doi.org/10.1121/1.4751541.

Gold, T. 1948. “Hearing. II. The Physical Basis of the Action of the Cochlea.” *Proceedings of the Royal Society of London B: Biological Sciences* 135 (881): 492–98. https://doi.org/10.1098/rspb.1948.0025.

Goldstein, J. L. 1967. “Auditory Nonlinearity.” *The Journal of the Acoustical Society of America* 41 (3): 676–99. https://doi.org/10.1121/1.1910396.

Hall, J. L. 1972a. “Auditory Distortion Products f_2-f_1 and $2f_1-f_2$.” *The Journal of the Acoustical Society of America* 51 (6B): 1863–71. https://doi.org/10.1121/1.1913045.

Hall, J. L. 1972b. “Monaural Phase Effect: Cancellation and Reinforcement of Distortion Products f_2-f_1 and $2f_1-f_2$.” *The Journal of the Acoustical Society of America* 51 (6B): 1872–81. https://doi.org/10.1121/1.1913046.

Hall, J. L. 1975. “Nonmonotonic Behavior of Distortion Product $2f_1-f_2$: Psychophysical Observations.” *The Journal of the Acoustical Society of America* 58 (5): 1046–50. https://doi.org/10.1121/1.380763.

Hall, James W., III. 2000. *Handbook of Otoacoustic Emissions*. San Diego, California: Singular Publishing Group.

Hällström, Gustav Gabriel. 1832. “Von den Combinationstönen.” *Annalen der Physik und Chemie* 100 (3): 438–66. https://doi.org/10.1002/andp.18321000303.

Harris, F. P., B. L. Lonsbury-Martin, B. B. Stagner, A. C. Coats, and G. K. Martin. 1989. “Acoustic Distortion Products in Humans: Systematic Changes in Amplitude as a Function of f_2/f_1 Ratio.” *The Journal of the Acoustical Society of America* 85 (1): 220–29. https://doi.org/10.1121/1.397728.

Harris, F.P., R. Probst, and L. Xu. 1992. “Suppression of the $2f_1-f_2$ Otoacoustic Emission in Humans.” *Hearing Research* 64 (1): 133–41. https://doi.org/10.1016/0378-5955(92)90175-M.

Hauser, R., and R. Probst. 1991. “The Influence of Systematic Primary-tone Level Variation L_2-L_1 on the Acoustic Distortion Product Emission $2f_1-f_2$ in Normal Human Ears.” *The Journal of the Acoustical Society of America* 89 (1): 280–86. https://doi.org/10.1121/1.400511.

Haworth, Christopher. 2011. “Composing with Absent Sound.” In *ICMC 2011*. University of Huddersfield, UK.

Haworth, Christopher. 2012. “Ear as Instrument.” *Leonardo Music Journal* 22 (22): 61–62. https://doi.org/10.1162/LMJ_a_00099.

Heller, Eric Johnson. 2013. *Why You Hear What You Hear: An Experiential Approach to Sound, Music, and Psychoacoustics*. Princeton: Princeton University Press.

Helmholtz, H. 1856. “Ueber Combinationstöne.” *Annalen der Physik und Chemie* 175 (12): 497–540. https://doi.org/10.1002/andp.18561751202.

Helmholtz, Hermann. 1954. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Translated by Alexander J. Ellis. Second English Edition. New York: Dover Publications.

Helmholtz, Hermann von. 1863. *Die Lehre von Den Tonempfindungen Als Physiologische Grundlage Für Die Theorie Der Musik*. Braunschweig: F. Vieweg und Sohn.

Hindemith, Paul. 1940. *Unterweisung Im Tonsatz 1. Theoretischer Teil*. Second Edition. Schott.

Hindemith, Paul. 1945. *Craft of Musical Composition*. Revised Edition. Vol. 1. New York: Associated Music Publishers, Inc.

Houtsma, A. J. M., and J. L. Goldstein. 1972. “The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition.” *The Journal of the Acoustical Society of America* 51 (2B): 520–29. https://doi.org/10.1121/1.1912873.

Humes, Larry E. 1979. “Perception of the Simple Difference Tone (f_2–f_1).” *The Journal of the Acoustical Society of America* 66 (4): 1064–74. https://doi.org/10.1121/1.383325.

Humes, Larry E. 1980a. “On the Nature of Two-tone Aural Nonlinearity.” *The Journal of the Acoustical Society of America* 67 (6): 2073–83. https://doi.org/10.1121/1.384451.

Humes, Larry E. 1980b. "Psychophysical Two-Tone Suppression as a Function of Input Level for $f_2/f_1 \gtrsim 1.0$." *The Journal of the Acoustical Society of America* 67 (5): 1759–63. https://doi.org/10.1121/1.384303.

Humes, Larry E. 1983. "Psychophysical Measures of Two-tone Suppression and Distortion Products ($2f_1-f_2$) and (f_2-f_1)." *The Journal of the Acoustical Society of America* 73 (3): 930–50. https://doi.org/10.1121/1.389018.

Humes, Larry E. 1985. "Cancellation Level and Phase of the (f_2-f_1) Distortion Product." *The Journal of the Acoustical Society of America* 78 (4): 1245–51. https://doi.org/10.1121/1.392893.

Janovsky, W. 1929. "Ueber Die Hörbarkeit von Verzerrungen (The Audibility of Distortion)." *E.N.T.*, no. 6: 421–39.

Jeffreys, H. 1961. *Theory of Probability*. Third Edition. Oxford, U.K.: Oxford University Press.

Kass, R. E., and A. E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90 (430): 773–95.

Kemp, D. T. 1978. "Stimulated Acoustic Emissions from within the Human Auditory System." *The Journal of the Acoustical Society of America* 64 (5): 1386–91.

Kemp, David T. 1998. "Otoacoustic Emissions: Distorted Echos of the Cochlea's Traveling Wave." In *Otoacoustic Emissions: Basic Science and Clinical Applications*, 1–59. San Diego and London: Singular Publishing Group.

Kemp, D. T. 2003. *The OAE Story: An Illustrated History of OAE Research and Applications through the First 25 Years*. Hatfield: Otodynamics.

Kemp, D. T. 2008. "Otoacoustic Emissions: Concepts and Origins." In *Active Processes and Otoacoustic Emissions in Hearing*, edited by Geoffrey A. Manley, Richard R. Fay, and Arthur N. Popper. Springer Handbook of Auditory Research. New York: Springer.

Kendall, Gary, Christopher Haworth, and Rodrigo F. Cadiz. 2012. "Sound Synthesis with Auditory Distortion Products." In *ICMC 2012*, 94–99. Ljubljana, Slovenia.

Kendall, Gary S., Christopher Haworth, and Rodrigo F. Cádiz. 2014. "Sound Synthesis with Auditory Distortion Products." *Computer Music Journal* 38 (4): 5–23. https://doi.org/10.1162/COMJ_a_00265.

Kirk, Jonathon. 2010. "Otoacoustic Emissions as a Compositional Tool." In *ICMC 2010*. New York.

König, Rudolph. 1876. "Ueber den Zusammenklang zweier Töne." *Annalen der Physik und Chemie* 233 (2): 177–237. https://doi.org/10.1002/andp.18762330202.

Krueger, F. 1900. "Beobachtungen an Zweiklängen." *Phil. Studien* 16: 307–79 and 568–664.

Krueger, F. 1910. "Die Theorie Der Konsonanz." *Psychol. Studien* 1 (1906), 2 (1907), 4 (1909), 5 (1910).

Kummer, Peter, Thomas Janssen, and Wolfgang Arnold. 1995. "Suppression Tuning Characteristics of the $2f_1{-}f_2$ Distortion Product Otoacoustic Emission in Humans." *The Journal of the Acoustical Society of America* 98 (1): 197–210. https://doi.org/10.1121/1.413747.

Lewis, D., and M. J. Larsen. 1937. “The Cancellation, Reinforcement and Measurement of Subjective Tones.” *Proceedings of the National Academy of Sciences* 23 (7): 415–21. https://doi.org/10.1073/pnas.23.7.415.

Licklider, J. C. R. 1951. “A Duplex Theory of Pitch Perception.” *Experientia* 7 (4): 128–34. https://doi.org/10.1007/BF02156143.

Marquardt, Torsten, Johannes Hensel, Dieter Mrowinski, and Gűnther Scholz. 2007. “Low-Frequency Characteristics of Human and Guinea Pig Cochleae.” *The Journal of the Acoustical Society of America* 121 (6): 3628. https://doi.org/10.1121/1.2722506.

Martin, Glen K, Eloy I Villasuso, Barden B Stagner, and Brenda L Lonsbury-Martin. 2003. “Suppression and Enhancement of Distortion-Product Otoacoustic Emissions by Interference Tones above F2. II. Findings in Humans.” *Hearing Research* 177 (1–2): 111–22. https://doi.org/10.1016/S0378-5955(03)00028-5.

Meenderink, S. W. F., and M. van der Heijden. 2010. “Reverse Cochlear Propagation in the Intact Cochlea of the Gerbil: Evidence for Slow Traveling Waves.” *Journal of Neurophysiology* 103 (3): 1448–55. https://doi.org/10.1152/jn.00899.2009.

Meenderink, Sebastiaan W. F., and Marcel van der Heijden. 2011. “Distortion Product Otoacoustic Emissions Evoked by Tone Complexes.” *Journal of the Association for Research in Otolaryngology* 12 (1): 29–44. https://doi.org/10.1007/s10162-010-0233-4.

Meyer, M. 1896. “Ueber Kombinationstöne Und Einige Hierzu in Beziehung Stehende Akustische Erscheinungen.” *Z. Psychol. Physiol. Sinnesorgane* 11.

Nielsen, Lars Holme, Gerald R. Popelka, Arne Nørby Rasmussen, and Poul Aabo Osterhammel. 1993. “Clinical Significance of Probe-Tone Frequency Ratio on Distortion Product Otoacoustic Emissions.” *Scandinavian Audiology* 22 (3): 159–64. https://doi.org/10.3109/01050399309047462.

Nuttall, Alfred L., Anthony J. Ricci, George Burwood, James M. Harte, Stefan Stenfelt, Per Cayé-Thomasen, Tianying Ren, et al. 2018. “A Mechanoelectrical Mechanism for Detection of Sound Envelopes in the Hearing Organ.” *Nature Communications* 9 (1): 4175. https://doi.org/10.1038/s41467-018-06725-w.

Ohm, G. S. 1839. “Bemerkungen Über Combinationstöne Und Stösse.” *Annalen Der Physik Und Chemie* 123 (7): 463–66. https://doi.org/10.1002/andp.18391230709.

Oliveros, Pauline. 1967. “Liner Notes for New Sounds in Electronic Music.” Odyssey.

Oliveros, Pauline. 2012. *Electronic Works 1965-1966*. Compact Disc. Paradigm Discs.

Orlarey, Yann, Stéphane Letz, and Dominique Fober. 2009. “Faust: An Efficient Functional Approach to DSP Programming.” In *New Computational Paradigms for Computer Music*. Paris, France: Delatour.

Oxenham, Andrew J., Christophe Micheyl, and Michael V. Keebler. 2009. “Can Temporal Fine Structure Represent the Fundamental Frequency of Unresolved Harmonics?” *The Journal of the Acoustical Society of America* 125 (4): 2189–99. https://doi.org/10.1121/1.3089220.

Pickles, James O. 2013. *An Introduction to the Physiology of Hearing*. Fourth edition. Leiden: Brill.

Plomp, R. 1965. “Detectability Threshold for Combination Tones.” *The Journal of the Acoustical Society of America* 37 (6): 1110–23. https://doi.org/10.1121/1.1909532.

Plomp, Reinier. 1966. *Experiments on Tone Perception*. Soesterberg: National Defense Research Organization TNO. Institute for Perception RVO-TNO.

Plomp, R. 1967. "Beats of Mistuned Consonances." *The Journal of the Acoustical Society of America* 42 (2): 462. https://doi.org/10.1121/1.1910602.

Plomp, Reinier. 1976. *Aspects of Tone Sensation A Psychophysical Study*. London, New York, San Francisco: Academic Press.

Popelka, Gerald R., Poul A. Osterhammel, Lars H. Nielsen, and Arne N. Rasmussen. 1993. "Growth of Distortion Product Otoacoustic Emissions with Primary-Tone Level in Humans." *Hearing Research* 71 (1–2): 12–22. https://doi.org/10.1016/0378-5955(93)90016-T.

Pressnitzer, D., and R.D. Patterson. 2001. "Distortion Products and the Perceived Pitch of Harmonic Complex Tones." In *Physiological and Psychophysical Bases of Auditory Function*, edited by D.J. Breebart, A.J.M. Houtsma, A. Kohlrausch, and V.F. Prijs. Maastricht, The Netherlands: Shaker Publishing BV.

Purcell, David W., Bernhard Ross, Terence W. Picton, and Christo Pantev. 2007. "Cortical Responses to the 2f1-F2 Combination Tone Measured Indirectly Using Magnetoencephalography." *The Journal of the Acoustical Society of America* 122 (2): 992–1003. https://doi.org/10.1121/1.2751250.

Raiffa, H., and R. Schlaifer. n.d. *Applied Statistical Decision Theory*. Cambridge, MA: MIT Press.

Rasmussen, Arne Nørby, Gerald R. Popelka, Poul Aabo Osterhammel, and Lars Holme Nielsen. 1993. "Clinical Significance of Relative Probe-Tone Levels on Distortion Product Otoacoustic Emissions." *Scandinavian Audiology* 22 (4): 223–29. https://doi.org/10.3109/01050399309047473.

Rayleigh (Strutt), J.W. 1896. *The Theory of Sound*. Second Edition. Vol. II. New York: Macmillan and Co.

Rickman, Maureen D., Mark E. Chertoff, and Kurt E. Hecox. 1991. "Electrophysiological Evidence of Nonlinear Distortion Products to Two-Tone Stimuli." *The Journal of the Acoustical Society of America* 89 (6): 2818–26. https://doi.org/10.1121/1.400720.

Romieu, J.B. 1751. *Nouvelle Découverte Des Sons Harmonique Graves Dont La Résonance Est Très Sensible Dans Les Accords Des Instruments à Vent*. Assemblée publique de la Société Royale des Sciences tenue dans la grande salle de l'Hôtel de Ville de Montpellier.

Schouten, J.F. 1938. "The Perception of Subjective Tones." In *K. Ned. Akad. Wet. Proc.*, 41:1086–93. Eindhoven, Holland.

Schouten, J.F. 1940. "The Perception of Pitch." *Philips Technical Review* 5 (10): 286–94.

Seebeck, A. 1841. "Beobachtungen über einige Bedingungen der Entstehung von Tönen." *Annalen der Physik und Chemie* 129 (7): 417–36.

Shannon, Robert V., and T. Houtgast. 1980. "Psychophysical Measurements Relating Suppression and Combination Tones." *The Journal of the Acoustical Society of America* 68 (3): 825–29. https://doi.org/10.1121/1.384821.

Sisto, Renata, Uzma Shaheen Wilson, Sumitrajit Dhar, and Arturo Moleti. 2018. "Modeling the Dependence of the Distortion Product Otoacoustic Emission Response on Primary Frequency Ratio." *Journal of the Association for Research in Otolaryngology* 19 (5): 511–22. https://doi.org/10.1007/s10162-018-0681-9.

Smoorenburg, Guido F. 1972. "Combination Tones and Their Origin." *The Journal of the Acoustical Society of America* 52 (2B): 615–32. https://doi.org/10.1121/1.1913152.

Sorge, G.A. 1744. *Anweisung Zur Stimmung Und Temperatur Sowohl Der Orgelwerke, Als Auch Anderer Instrumente, Sonderlich Aber Des Claviers*. Hamburg: Gedruckt mit Piscators schriften.

Stumpf, C. 1910. "Beobachtungen Über Kombinationstöne." *Z. Psychol* 55.

Tartini, G. 1754. *Trattato Di Musica Seconde La Vera Scienza Dell' Armonia*. Padova.

Vieth. 1805. "Ueber Combinationstöne, in Beziehung auf einige Streitschriften über sie zweier englischer Physiker, Th. Young und Jo. Gough." *Annalen der Physik* 21 (11): 265–314. https://doi.org/10.1002/andp.18050211102.

Warren, John M., and James P. Egan. 1951. "On the Accuracy of the Method of Best Beats for Determining the Intensity of a Tone." *The Journal of the Acoustical Society of America* 23 (1): 111–13. https://doi.org/10.1121/1.1906715.

Weber, R., and V. Mellert. 1975. "On the Nonmonotonic Behavior of Cubic Distortion Products in the Human Ear." *The Journal of the Acoustical Society of America* 57 (1): 207–14. https://doi.org/10.1121/1.380416.

Weber, Wilhelm. 1892. "Ueber die Tartini'schen Töne." In *Wilhelm Weber's Werke*, edited by Königlichen Gesellschaft der Wissenschaften, 360–64. Berlin, Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-662-24691-7_21.

Wegel, R. L., and C. E. Lane. 1924. “The Auditory Masking of One Pure Tone by Another and Its Probable Relation to the Dynamics of the Inner Ear.” *Physical Review* 23 (2): 266–85. https://doi.org/10.1103/PhysRev.23.266.

Wilson, J. P. 1980. “The Combination Tone, 2f1 – F2, in Psychophysics and Ear-Canal Recording.” In *Psychophysical, Physiological and Behavioural Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen, 43–52. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-9144-6_6.

Zurek, P.M., and R.M. Sachs. 1979. “Combination Tones at Frequencies Greater than the Primary Tones.” *Science* 205 (4406): 600–602.

Zwicker, E. 1955. “Der Ungewöhnliche Amplitudengang Der Nichtlinearen Verzerrungen Des Ohres.” *Acustica* 5: 67–74.

Zwicker, E. 1968. “Der Kubische Differenzton Und Die Erregung Des Gehors.” *Acustica* 20 (4): 206–9.

Zwicker, E. 1979a. “Different Behaviour of Quadratic and Cubic Difference Tones.” *Hearing Research* 1 (4): 283–92. https://doi.org/10.1016/0378-5955(79)90001-7.

Zwicker, E. 1979b. “Zur Nichtlinearität Ungerader Ordnung Des Gehörs.” *Acustica* 42: 149–57.

Zwicker, Eberhard. 1981. “Dependence of Level and Phase of the ($2f_1-f_2$) Cancellation Tone on Frequency Range, Frequency Difference, Level of Primaries, and Subject.” *The Journal of the Acoustical Society of America* 70 (5): 1277–88. https://doi.org/10.1121/1.387141.

Zwicker, Eberhard, and Frances P. Harris. 1990. "Psychoacoustical and Ear Canal Cancellation of ($2f_1-f_2$) Distortion Products." *The Journal of the Acoustical Society of America* 87 (6): 2583–91. https://doi.org/10.1121/1.399051.