2011

DELAY AND MODULATION PROCESSING AS MUSICAL TECHNIQUE IN ROCK

Mark Collins

DELAY AND MODULATION PROCESSING AS MUSICAL TECHNIQUE IN ROCK

(Thesis Format: Monograph)

By

Mark Collins

Graduate Program in Popular Music and Culture

A thesis submitted in partial fulfillment of the requirements
for the degree of:
Master of Arts

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
Dr. Jay Hodgson	Dr. Emily Abrams Ansari
Second Reader	Dr. Keir Keightley
Dr. Norma Coates	Dr. Norma Coates

The thesis by

Mark Collins

entitled:

Delay and Modulation Processing as Musical Technique in Rock

is accepted in partial fulfillment of the
requirements for the degree of
Master of Arts

Date______________________ ______________________________
Chair of the Thesis Examination Board

ABSTRACT

This thesis presents an analytic model for investigating the musical functions of delay and modulation signal processing in a pop/rock context. In so doing, it challenges prevalent academic assumptions about what, specifically, constitutes "musical practice," focusing analytic attention on musical procedures and terms reserved for recordists that, until very recently, have only registered in research as extra-musical technologizations of "live" exchange, if at all. Recordists do not *create* space via delay and modulation processing. Rather, they use delay and modulation processing, among other techniques, to provide psychoacoustic information which listeners require to infer space. Put differently, recordists use delay and modulation processing, among other techniques, to add psychoacoustic information to tracks and, in the process, to situate them within the broader space *represented* by a mix. This musical process is what I ultimately intend to elucidate through the model I present in this thesis.

KEYWORDS

Psychaoustics, Acoustics, Recording Practice, Recording, Recordist, Record Production, Signal Processing, Sound Source, Sound, Amplitude, Frequency, Spectrum, EQ, Wavelength, Sound Pressure Level, Sine Wave, Phase Shift, Phase, Masking, Hearing, Critical Bands, Pitch, Harmony, Reverb, Delay, Modulation, Modulated Delay, Compression, Analytic Model, Tracking, Mixing, Microphone, Preamplifier, ADC, DAC, Slapback, ADT, Multi-tap Delay, Echo, Flanging, Chorusing, Phasing, Tremolo, Amplitude Modulation, Ducked Reverb, Gated Reverb, Aesthetic, Musical Communication, Timbre, Audio, Popular Music, Popular Music Studies, Pop, Rock, Reggae, The Police, Bring On The Night

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Jay Hodgson, who has offered me invaluable guidance, perspective, encouragement, and support. His faith in my abilities, his willingness to help, and his pragmatism have steered me through the most daunting aspects of this project's evolution, and for that I can not thank him enough. I would also like to thank Dr. Norma Coates, who has been enthusiastic and supportive of my academic endeavours from the very start, and has stepped in time and time again in the roles of guide, sounding board, counsellor, and constructive critic. This thesis would not have been possible without Jay and Norma's guidance.

Thank-you to those on my defense committee, Dr. Keir Keightley, and Dr. Emily Abrams Ansari, and to all others who have helped along the way.

I am also infinitely grateful to my wife Natacha for her unconditional love and support on this long and life-changing road, and to my children Zoe and Zack for patience beyond their years and for their magical healing powers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

This thesis presents an analytic model for examining the musical functions of certain signal processing techniques, specifically, delay and modulation processing. In so doing, it challenges prevalent academic assumptions about what, specifically, constitutes "musical practice," focusing analytic attention on musical procedures and terms reserved for recordists that, until very recently, have only registered in research as extra-musical technologizations of "live" exchange, if at all (Hodgson 2010). Chapter One provides an academic context for this model, and explains its intended broader scholarly significance. Chapter Two elucidates the basic psychoacoustic assumptions on which the model is based, and explains how delay and modulation processing work. Chapter Three examines common ways that the psychoacoustic principles I examine in Chapter Two inhere in the production process, and, thus, it provides practical information analysts will require to apply the analytic model I present. Chapter Four presents the analytic model vis-a-vis a brief case study of The Police's "Bring On The Night" (1979), and in so doing tests its analytic viability. Chapter Five considers future research possibilities for generalizing, and expanding upon, my findings.

Recordists do not *create* space via delay and modulation processing. Rather, they use delay and modulation processing, among other techniques, to provide psychoacoustic information which listeners require to infer space. Put differently, recordists use delay and modulation processing, among other techniques, to add psychoacoustic information to tracks and, in the process, to situate them within the broader space *represented* by a mix. Vocals, for instance, cannot be physically positioned before, say, a drum-kit during playback; they can only be mixed to *seem* closer to the listener than the drums. Delay and modulation processing are two key tools recordists use to convey this impression, that is, to arrange sounds proximally in relation to listeners in a mix vis-a-vis psychoacoustic cues. This musical process is what I ultimately intend to elucidate through the model I present in this thesis.

The ability to arrange sound along a front-to-back "proximity plane," as Hodgson (2010) calls it, did not exist prior to the emergence of Recording Practice.[1] Before modern multitrack technology, composers simply lacked the technical means to achieve substantive movement along this plane. Though I'm certain experiments have been done in "live" practice to emulate proximity plane motion, the task of representing proximity plane motion in a musical context is reserved for recordists, insofar as it inheres in every recorded musical communication. Every recordist must determine musical uses for proximity plane motion, regardless of which genre they work in. For a record to achieve coherence, sounds must be arranged horizontally, vertically *and* proximally. Moreover, to achieve certain aesthetic ends, and to counteract certain technical limitations of the record medium, recordists often take it upon themselves to move sounds along the proximity plane multiple times over the course of the same arrangement, which is a physical impossibility in "live" performance.

---

[1] I use the capitalized title "Recording Practice" to refer to the practice of making and hearing music recordings as a complete communications system.

The proximity plane is as crucial and fundamental to recorded musical communications as is the horizontal plane, yet a specific analytic terminology for discussing its musical uses has yet to emerge. Analysts still discuss motion along the proximity plane in vague terms, often divorced from the techniques through which such motion is represented. Audio-engineering textbooks, for instance, typically fixate on technical specifications of technology, and they provide a psychoacoustic rationale for a few key uses, but they rarely address the aesthetic programs in whose service recordists create, and manipulate, the proximity plane of a mix. Musicological, historical and cultural studies of Recording Practice all tend to elide crucial technical details while they fixate on the (equally crucial) aesthetic and sociocultural ramifications of recorded musical communications, leaving informed readers to feel as though an entire analytic discourse has been created which fails to substantively address its primary subject, namely, Recording Practice *per se*.

This said, detailed and informative texts on Recording Practice have indeed emerged in the last two decades. Moreover, significant scholarly attention is now paid to the practice of making and hearing records from a multitude of disciplinary perspectives. Yet reading through this emerging literature is nonetheless an often frustrating endeavour, especially for those of us who do not simply analyze Recording Practice but, also, engage in it on a daily basis. Besides a few notable exceptions, the specific procedures recordists use to create recorded musical communications, let alone proximity plane motion, remain notably absent from the scholarly literature. Scholars most often study record reception when they ostensibly examine Recording Practice, or they study the way that records sound. Neither focus addresses Recording Practice *per se.* Recording Practice is the practice of *making* records, after all, not hearing them.

Of course, the present focus on reception in studies of Recording Practice has produced a fascinating record of the attitudes and cultural values which shape recorded musical communications. However, this (sometimes blinkered) analytic gaze tends to leave the recorded musical communication itself — the audio object, in all its polysemic slipperiness, which structures and enables reception in the first instance — almost completely transparent in analysis. I would not be so parochial as to suggest that reception is anything less than a fundamental and shaping component of musical meaning. I have, and continue, to learn much from reception-centric research. However, I do not believe it tells "the whole story," as it were; the narrative about musical meaning-making which analysis of reception alone construes is just not comprehensive. Receivers must receive *something* for reception to occur. No matter how polysemic the "text" receivers derive from a record, no matter how dialogic the utterance, as it were, they construe it from a limited set of terms furnished by recordists through record production.

Ultimately, I hope the analytic model I present in this thesis will contribute research on both the production *and* reception of recorded works, vis-a-vis focused analysis of delay and modulation processing and their psychoacoustic ramifications. I examine delay and modulation processing simply to test the analytic viability of my model, which details the broader musical ramifications of psychoacoustic cueing vis-a-vis signal processing; at a later date, I hope to examine the generalizability of my claims in regard to Recording Practice at large.

On a theoretical level, I hope this thesis will contribute to an analytic understanding of production and reception as related antipodes within a broader dialectic, namely, Recording Practice. One cannot receive a record unless it is first produced; musical communications remain silent collections of code which only *represent* sound without the interventions of reception (Hodgson 2010). As I hope to demonstrate through the literature review I offer below, research on Recording Practice remains deficient in analysis of record production, from which the dialectic of Recording Practice in the first instance emerges. I hope my thesis will ultimately provide interested researchers a few analytic tools to begin rectifying this imbalance.

Literature Review

Before I construct my model, I should first address a few ongoing academic debates about Recording Practice or, more specifically, how to study Recording Practice. At present, there exists a scholarly consensus that records are a primary means of communication between popular musicians and audiences. Moreover, most scholars agree that the expressive flow of a recorded musical communication, that is, the creation of emotional and musical tension by records, accrues via a complex interaction of literary, musical, and technological devices. However, the analytic tools that might enable scholars to effectively chart these interactions remain underdeveloped. The remainder of this chapter (headed "Literature Review") explores this lacuna through a survey of four common methodological approaches to the study of Recording Practice, and considers the strength and weaknesses of each for what I hope to achieve here. If I am to propose a new analytic approach to studying recorded musical communications, I shall first have to explain why, given the present academic context, methodological innovation is indeed required. To be clear, the following literature review is offered as a scholarly context for the model I present in this thesis, not as a definitive statement about the current state of research on Recording Practice at large.

Analysts usually follow one of four general methodological orthodoxies -- which, for the sake of convenience, I call "approaches" -- when discussing Recording Practice, namely, (i) a musicological approach, (ii) a cultural studies approach, (iii) a popular press approach, and (iv) a practical "how-to" approach. Of course, work in each of these categories is written for an ideal audience, that is, a "sympathetic" audience with a predisposition towards the paradigms and idiosyncrasies of the bodies of knowledge associated with each respective category. The strengths and weaknesses intrinsic to each approach largely depend on three crucial factors: (i) the audience receiving the information; (ii) the effectiveness of the devised framework for discussing the subject matter; and, most importantly, (iii) the ability to effectively discuss the musical elements that contribute to the musicality of the analyzed record.

My review of musicological frameworks for discussing records includes research by Albin Zak (2001), Nicholas Cook (2009), Theodore Gracyk (1996) and John Covach (1997). While by no means exhaustive, these studies comprise a reasonably comprehensive cross-section of the many ways that musicologists discuss records. I intend to demonstrate that musicological research on Recording Practice too often focuses on melody, harmony, and rhythm, and too often from the perspective of traditional music theory, or, it fixates on historically related extra-musicalities, leaving the particulars of recorded musical communications untouched.

Musical discussions in the realm of cultural studies tend to eschew direct discussion of musical particulars. Analysts of culture unsurprisingly opt to examine musical artifacts as cultural artifacts, requiring them to situate their work vis-a-vis sociology, anthropology, semiotics, political economy, poststructuralism, psychology, and other fields of cultural study which are not typically interested in musical technique *per se*. My analysis of this methodological approach focuses on Peter Doyle's *From 'My Blue Heaven' to 'Race with the Devil': echo, reverb and [dis]ordered space in early popular music recording* (2004). This section of the following literature review thus examines the strengths and weaknesses inherent in what I call "the cultural studies approaches" to analyzing Recording Practice, as represented by Doyle's work.

Writings about popular music in the popular press abound. These writings vary widely from publication to publication. They range from cultural studies, to romanticized aesthetic observations, to purely subjective opinions. I will not examine these accounts in detail in this literature review, because I believe that they present a range of disciplinary and methodological issues which are simply beyond the scope of this thesis to consider. It is sufficient to say that they do not exhibit engagement with the technical requirements of signal processing to be of use for constructing my model.

The final category I examine is what I call "the practical 'how-to' approach." This literature focuses squarely on the achievement of practical musical goals. Whether the topic is mixing audio or playing the guitar in a particular style, this writing focuses on providing straightforward practical instruction. In this category I include work by Roey Izhaki (2008), Bobby Owsinski (1999) and Alexander Case (2007).

To be clear, I am not interested in claiming that any one of these four "approaches" is preferable to another, except in the case of particular and situated analytic priorities. In this thesis I am interested in examining delay and modulation processing as musical endeavours, and, thus, as crucial components of recorded musical communications. I therefore value each of the approaches I examine below only for analytic insights they provide into the musical functions of delay and modulation processing. This is not to say that any of the approaches could not be re-tailored, or amended, to provide the technical information I claim they lack. In fact, readers will note that in constructing and presenting my model, I borrow freely from texts drawn from all four categories examined below.

The Musicological Approach[2]

Musicians, historians, acousticians, cultural theorists, performers, and everyone else, listen to and analyze records from their own personal analytic perspectives. Presumably, the audience a particular author belongs to, or is in dialog with, is the one s/he writes for. The contribution an author makes to a particular body of work is only one part of an ongoing discussion that has developed over time and through the work of others. Individuals unfamiliar with a particular body of knowledge will not necessarily agree with, or find useful, arguments and opinions deriving from another body of knowledge. While discussion of so-called "turf wars" within musicology is well beyond the scope of this paper, I do believe members of various analytic "factions" should

[2] The term 'musicology' has been defined in many different ways. As a method, it is a form of scholarship characterized by the procedures of research. A simple definition in these terms would be 'the scholarly study of music'. Traditionally, musicology has borrowed from 'art history for its historiographic paradigms and literary studies for its paleographic and philological principles' (Treitler, 1995). A committee of the American Musicological Society (AMS) in 1955 also defined musicology as 'a field of knowledge having as its object the investigation of the art of music as a physical, psychological, aesthetic, and cultural phenomenon.' The last of these four attributes gives the definition considerable breadth.

A third view, which neither of the definitions noted above fully implies, is based on the belief that the advanced study of music should be centered not just on music but also on musicians acting within a social and cultural environment. This shift from music as a product (which tends to imply fixity) to music as a process involving composer, performer and consumer (i.e. listeners) has involved new methods, some of them borrowed from the social sciences, particularly anthropology, ethnology, linguistics, sociology and more recently politics, gender studies and cultural theory. This type of inquiry is also associated with ethnomusicology. Harrison (1963) and other ethnomusicologists have suggested that 'It is the function of all musicology to be in fact ethnomusicology; that is, to take its range of research to include material that is termed "sociological" ... ' From *Grove Music Online* (www.oxfordmusiconline.com.proxy1.lib.uwo.ca:2048/subscriber/article/grove/music/46710pg1#S46710.1).

always remember that their particular branch of study is not supreme in the ongoing and comprehensive understanding of musical pursuits, but, rather, represents only one of many culturally and politically "interested," and institutionally situated, approaches to musical meaning-making.

Nicholas Cook's *Methods for Analysing Recordings* (2009) discusses the possible benefits that visual representations of music, generated by computer software, can bring to the study of a particular recording (e.g. spectrograms, Sonic Visualizer, "Soundbox" diagrams). Cook claims that these visual representations "heighten aural understanding of what is going on in the music," and that "[the] visualization represents something that is there to be heard in the music, but it adds something to the experience, refining and focusing your listening, and making you more aware of the sound space" (Cook 2009: 221, 223).

Offering visual cues to depict sonic events can indeed bring additional awareness to sonic aspects of a record which might otherwise remain unremarked. However, the types of images Cook suggests are static, and offer only a snapshot of just a few seconds of a production at a time. Music unfolds in real-time, and looking at static images that represent only several seconds of an entire recording does little good in determining anything about a recorded musical communication as a totality. For visual cues to be useful in the way that Cook proposes, they would have to unfold in real time *with* the music they purport to visually catalogue. Cultivating a tool of this sort is feasible, in my opinion, but would require that metadata about the mixing process be embedded in an audio file that would feed image-generating software, which would, in turn, require a massive paradigm shift from print to audiovisual emphasis in the presentation of analyses (i.e., analysts would have to focus on animating, rather than writing, their analyses).

Theodore Gracyk ties music to technology in *Rhythm and Noise: An Aesthetics of Rock*, but in a different way from Cook. Cook equates performance with the recorded work, that is, he discusses records as "recorded performances," while Gracyk describes popular music as a primarily recorded, rather than performed, communications paradigm. Gracyk argues that recordings themselves are the primary "art objects" in popular music. Thus, Gracyk continues, Recording Practice should occupy a foreground position in analysis of popular music.

For readers who are unaware of the arcane complexities of record production, Gracyk details common musical considerations which inhere in the production process. He elucidates not only the sound of records, but how those sounds were achieved. Gracyk warns that without considering the additional elements the recording process interjects over the course of a single recorded musical communication, a listening audience "can only respond at an unsophisticated level, confined to its [the recording/ song] basic features and obvious meanings. The work is reduced to its descriptive, narrative, or expressive elements" (Gracyk 1996: 45).

Gracyk's ultimate goal in presenting his approach is to establish rock records as complete works of art, and rock (popular music) as an autonomous art form with its own distinct aesthetic values and philosophies. He roots popular music squarely in the culture that creates it, namely, rock recordists and fans, and advocates for a closer examination of the production process in analytic assignments of musical meaning. In my opinion, Gracyk's main contribution to the discussion of rock records is that he establishes the production process itself as a shaping influence in the creation of recorded musical communications, not merely as a technologization of "live" exchange.

Published one year after *Rhythm and Noise*, John Covach's article *We Won't Get Fooled Again: Rock Music and Musical Analysis* (Covach 1997) would have benefitted from Gracyk's work on the importance of record production in popular music practice. Concerned mostly with establishing the need for music-theoretic modes in analysis of popular music, Covach focuses only on notatable aspects of communications of popular music, even as he warns readers that "as musical scholarship pays increasing attention to popular music, we need to be sure that we avoid falling into traps that silently reside within our own disciplines" (Covach 1997: 135).

Covach does not effectively make the connection between what is colloquially called "classical music" and popular music that his argument requires, in my opinion.[3] Covach does not demonstrate that rock and so-called "classical" music present sufficiently substantive overlap of musical values and practice to warrant the same analytic tools. In other words, though much of Covach's argument is convincing, he does

[3] I am aware of the current vogue for labelling what was once colloquially termed "classical" music, "art" music instead. I am extremely uncomfortable with the ramifications of this designation. I feel that categorizing so-called "classical" music as "art" music implies that other traditions are impoverished in some manner. I understand that the term "classical music" is considered to be inaccurate because it labels an entire galaxy of musical practice according to a single historical epoch, which historians now find an unduly domineering designation. I do not know what a solution to this issue might be, but I am convinced that the label "Art Music" is simply too problematic to reproduce, especially for those of us who consider popular music an equally valued "art music." That said, I will attempt to problematize my use of the term "classical" throughout, in deference to the present consensus.

not clarify which aspects of conventional theory should apply to rock, and which should be jettisoned. Moreover, much of rock music cannot be accurately notated using traditional music-theoretic tools, and the likes of David Brackett (1995) clarified long ago that notation imposes its own set of assumptions about what the musical experience can — and, indeed, *should* — be, privileging pitch relations, for instance, and harmonic design and metered rhythms, over timbre, dynamic contour and proximal location.

In *The Poetics of Rock: Cutting Tracks, Making Records* (Zak 2001), Albin Zak takes the discussion of the recorded work well beyond the level of detail Gracyk manages. Zak's method for discussing popular music entails a detailed account of the recording process complete with anecdotal information from popular musicians, recording engineers, and record producers. Zak approaches the discussion of sound by categorizing five phenomena that recordists manipulate: "1) performance, 2) timbre, 3) echo, 4) ambience (reverberation), and 5) texture" (Zak 2001: 49). He further breaks down texture, or "narrative dimension" as he calls it, into four categories: 1) width (stereo soundstage), 2) height (frequency spectrum), 3) depth (prominence), and 4) time. Many books on mixing methodology (discussed in more detail below) address the first three elements Zak examines. However, few address Zak's fourth dimension, namely, time, and those that do touch on it do not describe its function in aesthetic terms, as does Zak (2001: 160-161):

> The deliberate construction techniques used in fashioning the three dimensions represented by stereo, frequency range, and prominence are subsumed ultimately in temporal experience. That is, whatever constructive feats a mix entails, they are presented in the context of its fourth dimension, time. It is here, in their roles as participants in an unfolding sound drama, that the mix's various "shots" assume their contextual meaning. As noted earlier, a track's textural narrative has it own

> design whose unfolding may or may not coincide with that of the song. Even when its most obvious points of articulation do synchronize with those of the song, however, there are myriad textural variations that keep the mix fluid within structural subdivisions.

The strength of Zak's approach lies in its thorough engagement with the record making process, in my opinion. The breadth and scope of Zak's research goes beyond anything written academically for its time. Researchers in need of concrete information on the practice of signal processing might criticize Zak's work for a lack of technical detail, for instance. However, it is unreasonable to criticize the study for this reason, as such information is clearly beyond the scope of the work. Zak's contribution to musicology, and especially to popular music studies, is significant.

Musicology, a broad and diverse field, deals with all things musical, and integrates a variety of disciplines. The approaches addressed above represent a few of the many ways musicologists approach discussions of recorded sound in the context of popular music. While the sample I present is admittedly limited, it is indeed representative of the field at large. As the academic study of music expands its canon to include popular music texts, Recording Practice *per se* will need a more prominent role in research.

The Cultural Studies Approach

Cultural studies of records generally analyze Recording Practice as a social space, open to multiple and contested habitations, where producers and listeners negotiate identity or identities such as class, race, gender and sexuality. That is, cultural studies of records tend to focus primarily on the social ramifications of recorded musical communications, and they do not generally address Recording Practice *per se*. There are, however, a few notable exceptions, the most pertinent here being Peter Doyle's *From 'My Blue Heaven' to 'Race With The Devil': echo, reverb and [dis]ordered space in early popular music recording* (2004).

In *From 'My Blue Heaven' to 'Race With The Devil'*, Doyle (2004) historicizes the early use of reverberation as a musical figure in popular music productions. Doyle offers a brief, non-technical description of the difference between echo and reverb, and points out that "reverberation does much to define what we perceive as timbre, volume and sound colouration, and largely determines our perceptions of directionality and nearness" (Doyle 2004: 32). This description, while vague, at least refers to some psychoacoustic aspects of reverberation.

Through his comparison of so-called "spatial cue" creation in pre- and post-electric recording (i.e.., before and after roughly 1925, which marks the widespread adoption of electric transducers), Doyle (2004: 33-34) elucidates other ways recordings can create a kind of "pictorial spatiality" in the listener's mind. He states:

> Whereas pre-electric ... recordings ... had used sound effects - such as bird calls ... the sounding of foghorns - and elaborate verbal cues to construct a kind of pictorial spatiality, with electrical recording a real sense of spatial depth became possible ... A listener might now [referring to the increase in fidelity that came with electric

> recording] apprehend a recording and simultaneously experience a sense of a physical space, other than the actual space in which the playback device was located.

Through this description, Doyle distinguishes the use of verbal/aural cues to create *images* of space from the direct aural *experience* of space. Doyle attempts to musically contextualize reverb by discussing its use in several early productions, but his interpretation does not include musical technique *per se*.

Doyle's approach is erratic, in my opinion. In musical contexts, Doyle reduces reverb to a narrative effect relating only to lyrics. Each of his musical examples relies too heavily on lyrical content, or simply demonstrates reverb as a way of spatially foregrounding a primary 'voice' by moving other mix elements back along the proximity plane of a mix. This is simply too generalized an approach to be of use for my model. Doyle does discuss, say, microphone placement as a way of achieving a reverberant effect when tracking, but, even so, he does so in relation only to a lyrical narrative.

Doyle fixates on musical "meaning" in his analysis, a focus he shares with most analysts working from a similar institutional basis. While inquiries of this sort bring awareness and knowledge to different conceptions of musical practice, they offer little information about the music that these different conceptions produce. Cultural studies and musicology often disagree on what has primacy in popular music studies, but as Susan Fast (2000: 50) states, it is generally agreed that:

> ... musicologists need to be working with record producers and engineers to understand how technology is used to create certain effects, since the sound (i.e. timbre, production values, etc.) of many recordings is critically important both to the musicians and to those who listen to the music; we also need to know more about how the instruments used in popular music work ... and how the various sounds are produced on them.

Interestingly, Fast proposes this so that cultural meanings can be ascribed to music. Regardless of the motivation, the recognition that those involved in record production should be included in the scholarly study of popular music is a welcome notion.

The "How-To" (or, Practical) Approach

Authors working in what I call "the 'how-to' market" assume that their readers have the desire to acquire a certain skill set, or aspire to understand the detailed processes professionals engage in when *producing* records. Presumably, the majority of readers that study a practical guide look to apply the information they glean toward their own endeavours. This is not to say that literature from the other categories requires no effort to engage with. On the contrary, some of the concepts engaged by musicologists, cultural theorists and journalists require years of study to adequately grasp. However, the literature from the "how-to" market mandates the direct *application* of knowledge (e.g. exercises), and active *participation* from the reader (e.g. listening tests, etc.). This is, in my opinion, a crucial difference between this category and those that I have already examined.[4]

Roey Izhaki's *Mixing Audio: Concepts, Practices and Tools* (2008) is precisely the kind of guide that requires reader interaction. Izhaki's book deals exclusively with the art of mixing audio in a record making context. He discusses sound directly, and

[4] Much of the literature in this category also has a multimedia component. I find it important to bring attention to this fact as this is perhaps the only category of literature that supplies the reader with sound files/music for demonstrative and comparative purposes. Academic journals and popular press do not exploit the use of multimedia content in this way.

offers the reader listening examples and mixing exercises on the accompanying DVD. *Mixing Audio* is organized into three basic categories: concepts and practices, tools, and sample mixes. The 'concepts and practices' category deals with: mixing philosophy, axioms of mixing, mixing processes, common issues encountered when mixing, and mixing domains and objectives. Each of these areas deals with either conceptualizing a mix, that is, how disparate sonic elements come together to form a whole, or how a particular facet of mixing directly affects the sound of an overall mix. The 'tools' section discusses the tools at a mixing engineer's disposal (e.g. monitors, consoles, panning, eq, compressors, etc.), and their detailed function (how they change sound, and why they are typically put to use in mixing situations). Sound examples and diagrams complement the text throughout. The third and final section, 'sample mixes,' details multiple mix approaches for four different songs from four different genres, and offers raw tracks (or stems) for the reader to mix (all included on the DVD). This section details genre-specific mixing techniques, and then applies them in different mix contexts (with detailed mix notes for each multi-track) so the reader can *hear* the application (in context) of the principles and techniques outlined throughout the book.

Izhaki approaches the discussion of sound in *Mixing Audio* in a practical and organized way. The topic itself, namely, mixing, is an integral component of the record production process. For example, Izhaki categorizes mix elements (kick drum, guitars, etc.) according to how they must be dealt with because of their masking qualities, and their function (interest, pulse/groove, harmonic support, etc.) within a mix. Once categorized, Izhaki evaluates these sounds against their sonic domain properties (frequency, amplitude, stereo, depth) for placement and functional effectiveness. Breaking sonic components into functional and relational categories offers a way of discussing sound based on its reception in a musical context. Izhaki's approach allows for the creation of a discussion platform that helps describe sound's narrative function in musical terms.

By detailing the shaping influence specific processing tools have on sound (e.g. the effect compression has on the dynamic contour of a sound), and providing their typical musical context, Izhaki describes how the tools of record production directly shape recorded sound. This discussion elucidates generally understudied aspects of record production, and presents a discussion framework for addressing recordings of popular music in a way that limits the need for subjective commentary.

Bobby Owsinski, in *The Mixing Engineer's Handbook*, takes a slightly different approach in his explanation of the mixing process. Far more colloquial in tone, Owsinski covers fewer concepts than Izhaki, and with much less technical detail and precision. For example, Owsinski reduces discussion of the frequency domain to broad equalization techniques. Owsinski attempts to compensate for simplified explanations by including an abundance of anecdotes from famous mix engineers about their approaches to mixing. Many of the anecdotes describe how each respective individual 'feels' or 'hears' their way through a particular mixing technique or problem. For example, Ed Seay, discussing vocal equalization, states, "On a vocal sometimes I think, 'Does this vocal need a diet plan? Does he need to lose some flab down there?' Or sometimes, 'We need some weight on this guy so let's add some 300 cycles and make him sound a little more important'" (Owsinski 1999: 34). *The Mixing Engineer's Handbook* reads more like a 'tips and tricks' book than a manual detailing the practical vagaries of mixing. The methodology lacks formal organization, consistent language when referring to the same things, and feels somewhat disconnected as the flow jumps from one individual's experience to the next. Due to the lack of precision and organization, readers are left to interpret the anecdotal information as they see fit.

In contrast to Owsinksi's lack of organization and technical depth, Alexander Case, in his book *Sound F/X: Unlocking the Creative Potential of Recording Studio Effects* (2007), systematically examines every species of audio signal processing in great technical detail. Case's inclusion of acoustic engineering formulae and psychoacoustic theory provide a complete and thorough study of signal processing. While not as attentive to the mixing process as Izhaki, which is not surprising given that his topic is signal processing *per se*, Case does offer a basic guide that addresses signal processing applications in a mixing context. For each kind of signal processing (e.g. compression, delay, reverb, etc.) Case includes a selected discography for the reader to reference.

Case's methodology for discussing sound relies on: (i) a systematic breakdown of sonic parameters (pitch, amplitude envelope [attack, decay, sustain, release - relating to timbre], rhythm, spectral content, etc.) to differentiate audible characteristics of a sound-source; (ii) a discussion of signal processing techniques and their effect on sound and listener reception; (iii) standard uses for studio effects in production situations; (iv) guided listening examples; (v) psychoacoustic theory; and (vi) the physical characteristics of sound as per acoustic engineering principles. The technical depth of the discussion makes the intended audience for *Sound F/X* difficult to categorically identify. Case's effort draws on expertise from well established academic disciplines, and appeals to practitioners as well. This puts *Sound F/X* in a unique position amongst all the literature discussed in this review.

"How-to" research is obviously focused on practice, just as cultural studies are focused on culture, and musicological analyses are focused on what the discipline defines as "music." Unlike any of the other approaches, however, this method of discussion is drawn from within the world of record production, which has its benefits and its dangers (i.e.., simply because they make music does not necessarily make rock musicians experts on how their work should be academically analyzed). Practical information does offer a way to understand record-making as recordists conceive it, however. Work within this category certainly requires further consideration by academics hoping to engage in discussions of Recording Practice.

SUMMARY

In this literature review, I have examined four prevalent approaches to the analysis of recorded musical communications, with an eye towards what they can contribute to the analytic model I will now construct. What I have called "the musicological approach" provides an aesthetic context for technical detail, but ultimately fails to address Recording Practice as an agglomeration of musical techniques, which my model requires. What I call "the cultural studies approach" either ignores musical technique altogether, or, at best, discusses it in an unsophisticated manner; and, thus, it will bear limited fruit for the purposes of this inquiry. The “popular press” or “trade” approach is no help; its authors remain overly fixated on their own subjective responses to music to provide the technical detail I require. Finally, the practical guides I examine require broadening through the aesthetic considerations of musicology and cultural studies, though they are, indeed, most useful. A new methodology for studying popular music is ultimately required if analysts are to engage with the particulars of recorded musical communications in a meaningful, substantive manner. It is my ultimate hope that the analytic model I now turn my attention to constructing will contribute to the creation of such a methodology.

CHAPTER TWO

PSYCHOACOUSTICS: ASSUMPTIONS FOR THE ANALYTIC MODEL

In this chapter, I elucidate the basic psychoacoustic assumptions on which the broader analytic model I present in this thesis is based, and I situate a basic explanation of how delay and modulation processing work vis-a-vis those assumptions. In the interests of clarity, I have opted to divide this chapter into four parts. Parts One and Two, headed "Sound" and "Hearing," respectively, explain how psychoacoustics conceives hearing. According to psychoacoustic theory, there are two components to sound, specifically, (i) the actual physical matter of sound and (ii) the physiological responses it provokes in the human hearing mechanism, a process called "hearing." Parts Three and Four concern the behavior of sound in space, and the way that delay and modulation processing mimics that behavior.

Admittedly, psychoacoustics does not in any way account for cultural conditioning, which inheres in hearing. After all, sound doesn't simply physiologically register without the commentary human consciousness automatically provides for what registers through the senses; it is, indeed, reasoned and interpreted by the hearer, according to how they have been socialized to do so. For instance, a Western listener confronted with Carnatic or Hindustani music might interpret the prevalent quarter-tones as "noise." However, psychoacoustics remains uninterested in such interpretations. The facets of hearing which psychoacoustics theorizes occur well before any broader, cultural interpretations accrue. Psychoacoustics is, indeed, only interested in (i) the physics of sound and (ii) the physiology of hearing. It is interested, in other words, with elucidating the materials from which conscious interpretations and meanings are forged.

To be clear, my model will explain how psychoacoustics is, indeed, interested in explaining meaning making, but on a very fundamental level. The "meanings" I determine concern spatialization and acoustics, and the way that listeners interpret these from the psychoacoustic cues furnished by recordists through delay and modulation processing. Clearly, the broader interpretations forged by those same listeners are vast and require extensive research all their own. As I noted in the introduction, my primary aim in constructing this model is to interject a degree of specificity in understanding the raw materials out of which receivers fashion their receptions, which are furnished through production practices like signal processing. I do not intend for this model to be used as a justification for ignoring the crucial role of reception in musical communication, as I fear some readers may choose to do (this is undoubtedly clear from the number of these caveats). Nonetheless, a base understanding of psychoacoustics is required to understand the contributions my analytic model offers. I now turn my attention to providing that base knowledge in this chapter. The next chapter will survey "common practice" musical applications of the concepts I survey in this chapter.

PART ONE: SOUND

Sound mechanically disturbs the medium through which it travels. Vibrations produced by that disturbance, called a "sound source," agitate the molecules immediately surrounding it, forcing them to compress and rarefy in recurring patterns called "sound waves" (see Figure 2.01).[5] These waves propagate longitudinally, that is, they compress and rarefy outward from the vibrating sound source, in numerous directions at once. In air, which is the usual propagating medium for music, the disturbance moves so quickly that no heat transfer occurs from compression to rarefaction.

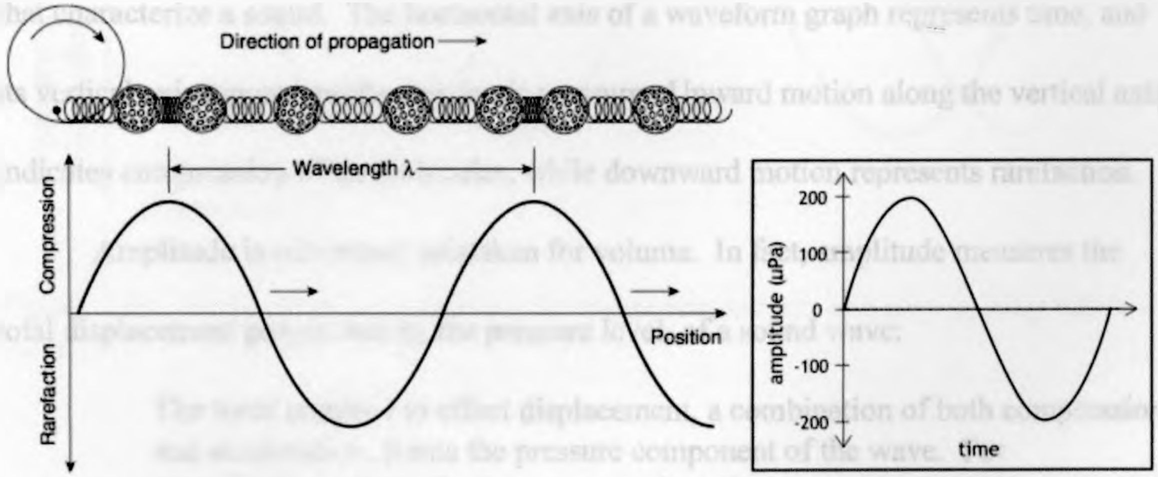


Figure 2.01. A propagating sine wave.

[5] While sound does indeed travel through solids and liquids the only medium under discussion in this study is air.

"Frequency" and "wavelength" measure the pressure variances within a sound wave as functions of time and space. A sine wave represents the simplest form of periodic excitation. Sine waves have three parameters: amplitude, frequency (and, relatedly, wavelength), and phase.[6] I explain each of these components in turn below.

Amplitude, pt. 1

Amplitude refers to the total displacement of air molecules that a sound wave generates, providing the characteristically erratic up-and-down motion of waveform charts. Usually represented in graph form, waveforms depict the changes in air pressure that characterize a sound. The horizontal axis of a waveform graph represents time, and its vertical axis represents changes in air pressure. Upward motion along the vertical axis indicates compression of air molecules, while downward motion represents rarefaction.

Amplitude is commonly mistaken for volume. In fact, amplitude measures the total displacement power, that is, the pressure level, of a sound wave:

> The force required to effect displacement, a combination of both compression and acceleration, forms the pressure component of the wave. For compression and rarefaction to occur air molecules must move closer together or further apart. Movement implies velocity, so there must be a velocity component which is associated with the displacement component of the sound wave ... The velocity component reaches its peak between the compressions and rarefactions, and for a sine wave displacement component the associated velocity component is a cosine [see Figure 2.02] ... The force required to accelerate molecules forms the pressure component of the wave. This is associated with the velocity component of the propagating wave and therefore is in phase with it. That is, if the velocity component is a cosine then the pressure component will also be a cosine. Thus, a sound wave has

---

[6] Sine waves do not occur in nature. However, the behaviour of more complex wave forms can always be described in terms of sine waves (Howard and Angus 2006: 52).

both pressure and velocity components that travel through [air] at the same speed (Howard and Angus 2006: 12-13).[7]

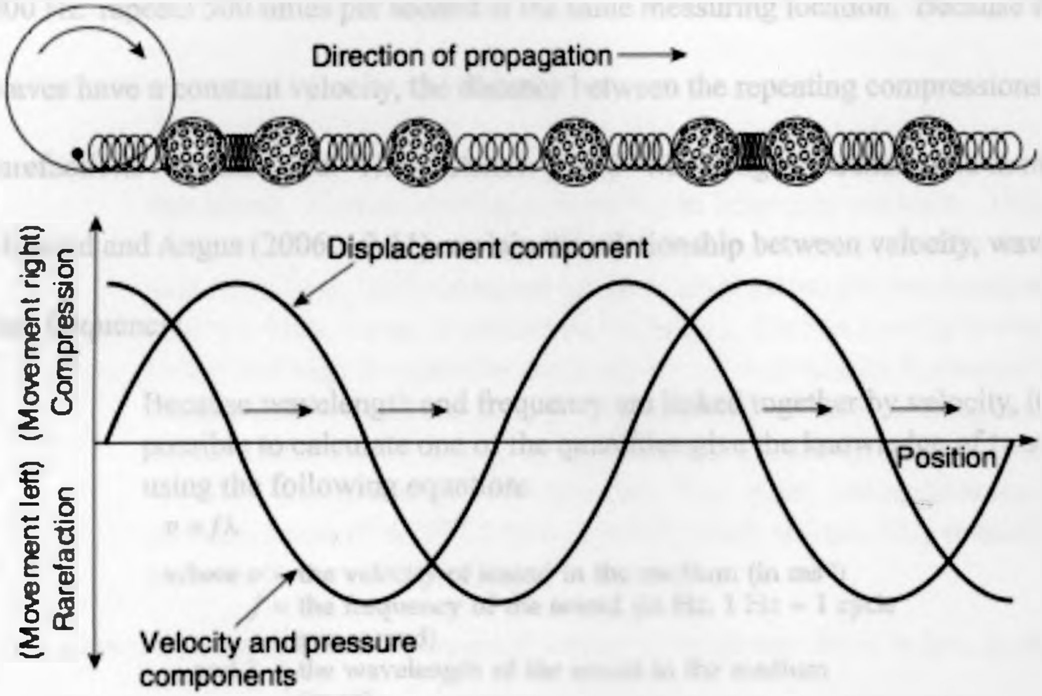


Figure 2.02. Pressure, velocity, and displacement components of a propagating sine wave.

[7] Sound pressure level (SPL) is a more accurate measurement of perceived loudness. This will be covered in more detail later in this chapter.

Frequency

Frequency, expressed in Hertz (Hz) or kiloHertz (kHz), identifies the number of times a wave passes through a given point per second. A sine wave with a frequency of 500 Hz repeats 500 times per second at the same measuring location. Because these waves have a constant velocity, the distance between the repeating compressions and rarefactions remains fixed. This distance, called "wavelength," is measured in meters. Howard and Angus (2006: 10-11) explain the relationship between velocity, wavelength, and frequency:

> Because wavelength and frequency are linked together by velocity, it is possible to calculate one of the quantities give the knowledge of two others using the following equation:
>
> $$v = f\lambda$$
>
> where v = the velocity of sound in the medium (in ms^{-1})
> f = the frequency of the sound (in Hz, 1 Hz = 1 cycle per second)
> and λ = the wavelength of the sound in the medium (in m)
>
> This equation can be used to calculate the frequency given the wavelength, the wavelength given the frequency, and even the speed of sound in [air] given the frequency and wavelength ... Calculations for the wavelength of sound, being propagated in air at 20°C, at 20 Hz and 20 kHz:
>
> $$\lambda = \frac{v}{f}$$
>
> which gives:
>
> $$\lambda = \frac{344}{20} = 17.2 \text{ m for 20 Hz}$$
>
> and
>
> $$\lambda = \frac{344}{20 \times 10^{3}} = 1.72 \text{ cm for 20 kHz}$$

As seen above, higher frequencies generate shorter wavelengths, while lower frequencies generate longer ones. Humans perceive frequency, and wavelength, as pitch. Because of this, an increase in a soundwave's Hertz value, that is, an increase in the rapidity of the sound wave's recurrence, causes its perceived pitch to rise. A frequency range of 20 Hz to 20 kHz establishes the lower and upper limits of human hearing in which such changes may occur, for the ear actively rejects frequencies below roughly 20Hz and above about 20kHz. Alexander Case (2007: 73-74) explains the relationship between frequency and perceived loudness:

> Human sensitivity to sound is strongly governed by the frequency content of that sound. Human hearing is never flat in frequency response. Uniform loudness comes from a range of sound pressure levels across the frequency axis. Similarly, uniform sound pressure level across the frequency axis would lead to a broad range of perceived loudness ... Human hearing is less sensitive to low and high frequencies and most sensitive to middle frequencies. this trend is shown to exist at a range of amplitudes. Human hearing consistently remains most sensitive to [the] upper-middle frequency range near about 3500 Hz, across a range of amplitudes - from quiet, just-audible sound pressure levels (0 dBSPL) up to painfully loud, and possibly unhealthy sound pressure levels (100 dBSPL).

I examine the phenomenon of "perceived loudness" in greater detail below, in the section headed "Sound Pressure Level." Before I can explain this phenomenon further, however, I must first explain what is meant by the term "phase."

Phase

Phase refers to a periodic waveform's current position with respect to its full cycle. Each of these cycles is measured in degrees, and phase *shift* occurs when two waves of the same frequency arrive at the same physical location at different points in their cycles (see Figure 2.03). Because recorded music contains many frequencies, the waveforms that make up the individual tracks of a multitrack recording are necessarily complex. Delaying a copy of a complex waveform will result in each frequency within the waveform having its own degree of phase shift against the original track. Roey Izhaki (2008: 165) discusses phase relations amongst waveforms:

> [W]e only consider phase in relation to similar waveforms. We have to define similar first:
> ***Identical waveforms*** - ... , these are similar in every way and usually the outcome of duplication. For example, by duplicating a vocal track in an audio sequencer we get two identical waveforms ...
> ***Waveforms of the same event*** - two microphones capturing the same musical event. For example, two microphones fronting an acoustic guitar.

When recordists align similar soundwaves so they are slightly out of phase, they introduce so-called "constructive" and "destructive" interferences. Some frequencies double in amplitude, while others cancel altogether, and this causes the tonal quality of the summed output to change. This process is known as comb-filtering, and Alexander Case (2007: 232) explains the derivation of the term:

> Combining a musical waveform with a delayed version of itself radically modifies the frequency content of the signal ... The intermediate frequencies experience something in between outright cancellation and full on doubling ... Taking a complex sound like guitar, which has sound energy at a range of different frequencies, and mixing in a delayed version of itself at the same amplitude, will cut certain frequencies and boost others. This is called comb filtering, because the alteration in the frequency content of the signal looks like the teeth on a comb.[8]

---

[8] The interplay between the original and delayed signal causes constructive and destructive and destructive interference. The frequency response of a comb-filter consists of a series of regularly spaced spikes and dips, giving the appearance of a comb (see Figure 2.04).

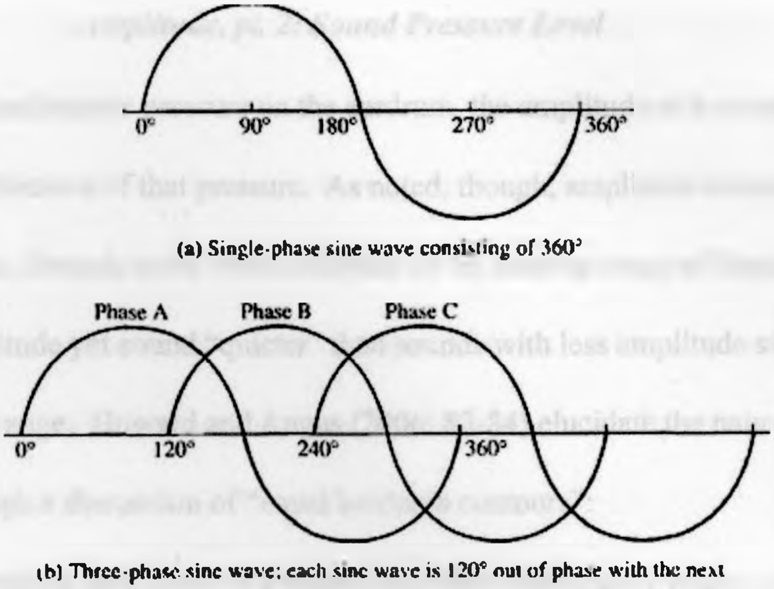


Figure 2.03. Phase shift.

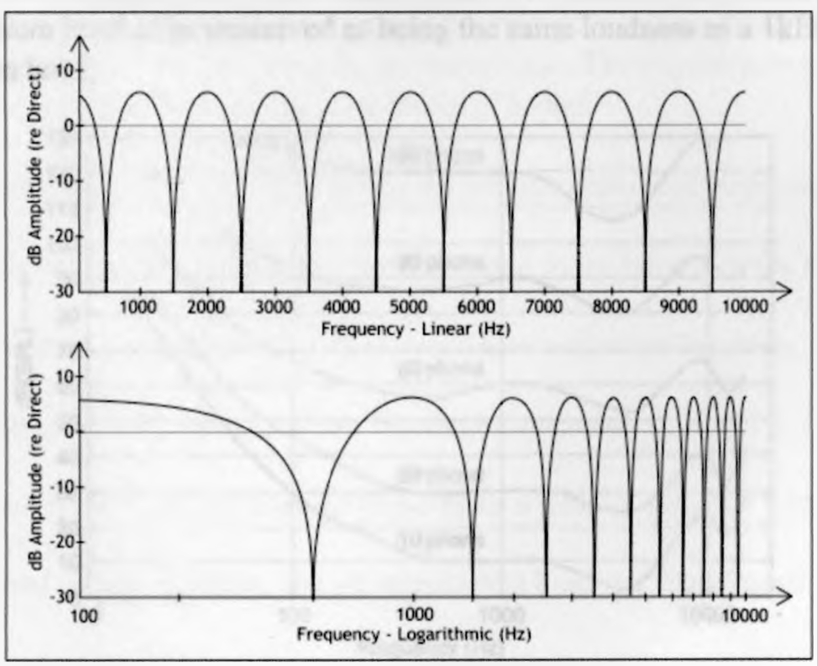
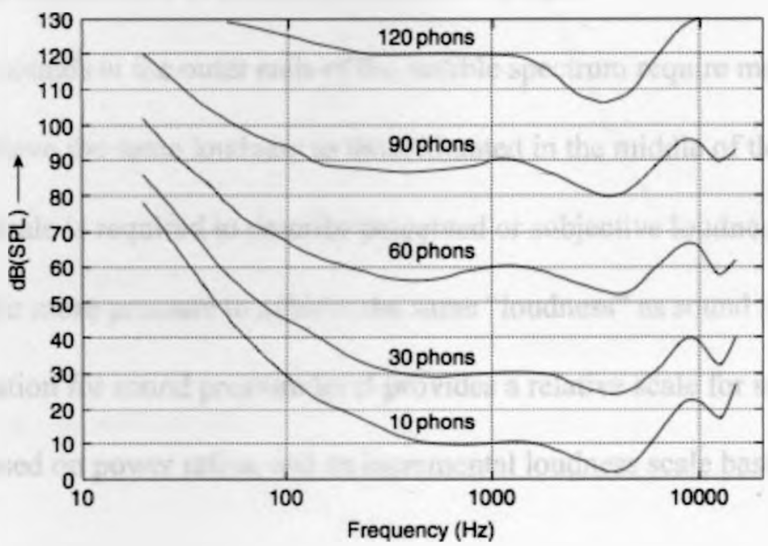


Figure 2.04. Comb filter.

Amplitude, pt. 2: Sound Pressure Level

Because sound exerts pressure on the eardrum, the amplitude of a sound wave can be measured as a function of that pressure. As noted, though, amplitude does not indicate *perceived* loudness. Sounds at the outer extremes of the hearing range of frequencies can have a larger amplitude yet sound "quieter" than sounds with less amplitude situated in the middle of that range. Howard and Angus (2006: 83-84) elucidate the nature of this phenomenon through a discussion of "equal loudness contours":

> The pressure amplitude of a sound wave does not directly relate to its perceived loudness ... How can this be so? The answer is that the sounds are at different frequencies and the sensitivity of our hearing varies as the frequency varies. Figure 2.12 [included below] shows the equal loudness contours for the human ear. These contours, originally measured by Fletcher and Munson (1933) and by others since, represent the relationship between the measured sound pressure level and the perceived loudness of the sound. The curves show how loud a sound must be in terms of the measured sound pressure level to be perceived as being the same loudness as a 1kHz tone of a given level.



Pressure *levels* created by a sound source are in constant flux. Therefore, pressure level — or, perceived loudness — is measured by calculating the root mean square (rms) pressure of a sound wave. Sound sources can vary over a range of pressure amplitudes greater than a million to one, and because of the way the human hearing mechanism registers sound, a logarithmic scale expresses sound pressure level most accurately. Howard and Angus (2006: 17-18) discuss the measurement of sound pressure at the threshold of hearing of a standard frequency:

> This scale is based on the ratio of the actual sound pressure to the notional threshold of hearing at 1kHz of 20 μPa. Thus the sound pressure level (SPL) is defined as:
>
> $$SPL = 20 \log_{10}\left(\frac{p_{actual}}{p_{ref}}\right)$$
>
> where p_{actual} = the actual pressure level (in Pa)
> and p_{ref} = the reference pressure level (20 μPa)
>
> The multiplier of 20 has a twofold purpose. The first is to make the result a number in which an integer change is approximately equal to the smallest change that can be perceived by the human ear. The second is to provide some equivalence to intensity measures of sound.

Because sounds at the outer ends of the audible spectrum require more amplitude (pressure) to achieve the same loudness as those situated in the middle of the spectrum, an equivalency scale is required to describe perceived or subjective loudness. Sound at 20 Hz requires far more pressure to achieve the same "loudness" as sound at 1 kHz. The logarithmic equation for sound pressure level provides a relative scale for sound measurement based on power ratios, and an incremental loudness scale based on *perceivable* differences in pressure. The decibel (dB) is the most common unit of measurement used to express this.[9]

Often confused with amplitude, loudness perception (colloquially called "volume") is a result of the brain calculating peak amplitude and the amount of time a waveform spends at, or near, peak amplitude (Hodgson 2010: 6). The human brain calculates the average pressure of a waveform, and humans perceive these averages as variations in loudness. The difference between a waveform's peak and rms amplitudes is called "crest factor." Crest factor refers to the difference between a signal's peak and rms amplitudes (see Figure 2.05).

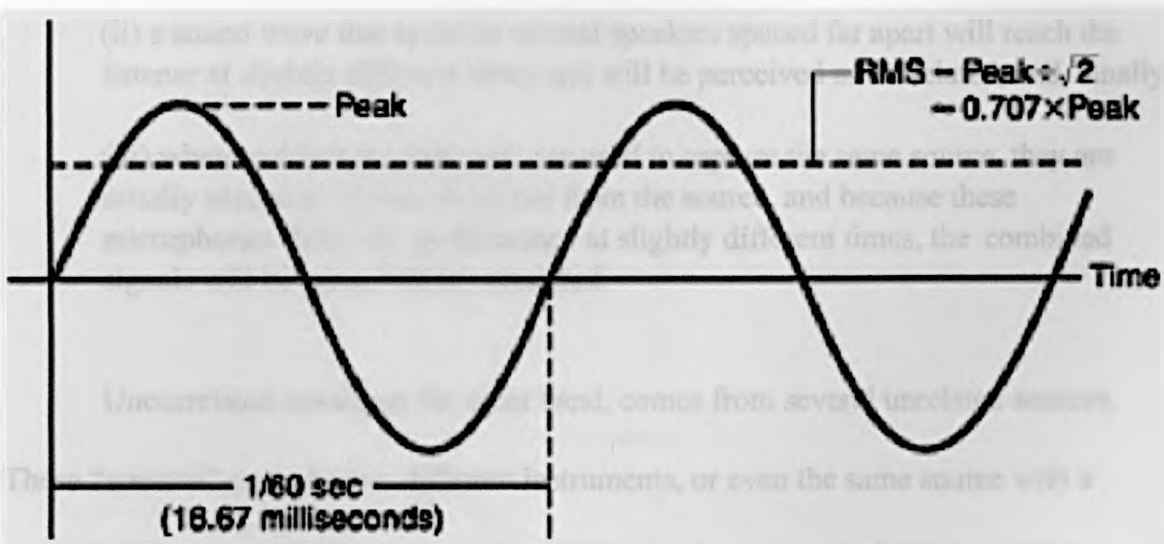


Figure 2.05. Peak and RMS levels.

---

[9] "The 'bel' is named after Alexander Graham Bell, and the number of bels is defined as the common logarithm of two powers. Thus, two powers, one of which is ten times the other, will differ by 1 bel ... Decibels have caused untold confusion among audio people, and most of this is due to the failure to realize that decibels are not quantities of anything and can represent only power ratios." (White and Louie, 2005: 99-100)

Correlated and Uncorrelated Sound Sources

Thus far, I have used only sine waves to demonstrate the physical behaviour of sound. However, listeners rarely encounter such a sound source for an extended period of time. In most productions, regardless of their intended market(s), multiple sound sources compete for spectral prominence in the mix.

Pyschoacousticians refer to sound that emerges from several related sources at once as "correlated" sound (Howard & Angus 2006: 20). Correlated sound can be produced in a number of ways. Representative examples include:

> (i) reflections (delayed sound) from nearby surfaces (walls, floor, etc.): if the delay is short, the reflection is similar enough to the original that receivers perceive it as "correlated" with the source;
>
> (ii) a sound wave that is fed to several speakers spaced far apart will reach the listener at slightly different times and will be perceived as correlated; and, finally,
>
> (iii) when multiple microphones are used to capture the same source, they are usually placed at various distances from the source, and because these microphones 'hear' the performance at slightly different times, the combined signals will be perceived as correlated.

Uncorrelated sound, on the other hand, comes from several unrelated sources. These "sources" could be two different instruments, or even the same source with a considerable amount of delay caused by reflections (or, particular delay processing techniques designed to emulate reflections). Even though they come from a single source, recordists consider extremely uncorrelated sounds to exhibit different waveforms. The pitch, amplitude and wave-shape of the trailing waveform — most obviously, the high-frequency content — changes, and, as such, appears unrelated to the original waveform.

Recordists combine complex waveforms when making recorded musical communications, and the resulting sound levels depend on whether these waveforms are correlated or uncorrelated. The phase relationship between correlated sources determines the overall increase or decrease in amplitude of combined signals. For example, if correlated sounds arrive completely in-phase they will double in amplitude, but if they arrive in a state of anti-phase they cancel entirely (also referred to as "nulling"; see Figure 2.06). In most cases, however, correlated sounds arrive at the listener's position with a phase relationship somewhere between these two extremes, and this increases some of the frequencies while canceling others, all of which alters the perceived timbre of the original sound source, and induces the tonal distortion colloquially called "comb filtering" by recordists and psychoacousticians, which I discussed earlier.

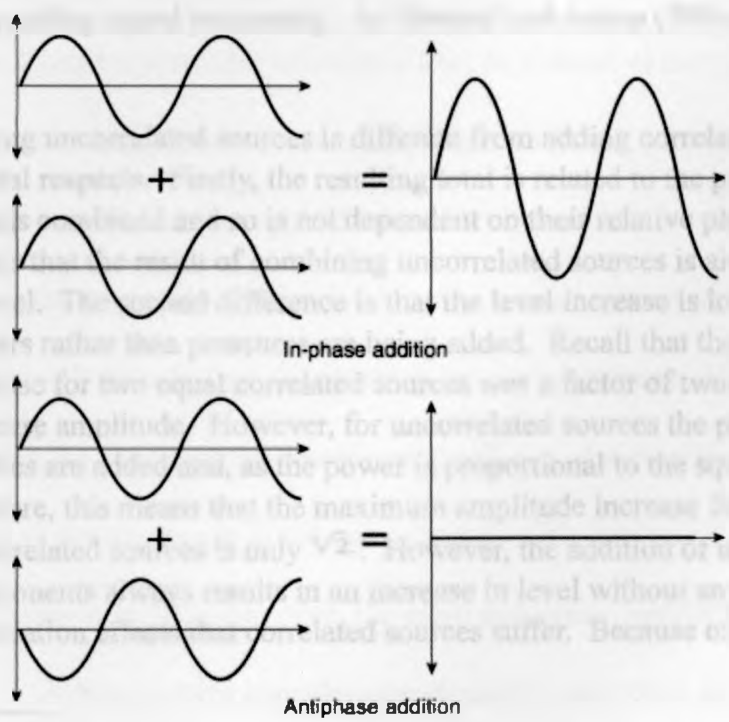


Figure 2.06. Addition of sine waves of different phases.

Combination of uncorrelated sources does not alter phase relationships in the same way, for these sources do not cause frequency cancellations to occur or pressure levels to increase. Nonetheless, combining uncorrelated sources does increase overall output level at the master bus (multitrack recordings reduce to stereo Left + Right), because sound *power* levels are summed instead of sound *pressure* levels.[10] For example, bussing combines several non-correlated sounds together. Signals sent to the bus may not show clipping at the individual level, but the summed signals can cause the bus channel to overload.

Recordists must consider the effects of combining correlated and uncorrelated sounds on a regular basis, for at each stage of production they must anticipate the resulting tonal quality and amplitude in order to make appropriate equipment selection and decisions regarding signal processing. As Howard and Angus (2006: 24-25) explain:

> Adding uncorrelated sources is different from adding correlated sources in several respects. Firstly, the resulting total is related to the power of the signals combined and so is not dependent on their relative phases. This means that the result of combining uncorrelated sources is always an increase in level. The second difference is that the level increase is lower because powers rather than pressures are being added. Recall that the maximum increase for two equal correlated sources was a factor of two increases in pressure amplitude. However, for uncorrelated sources the powers of the sources are added and, as the power is proportional to the square of the pressure, this means that the maximum amplitude increase for two uncorrelated sources is only $\sqrt{2}$. However, the addition of uncorrelated components always results in an increase in level without any of the cancelation effects that correlated sources suffer. Because of the lack of

[10] Adding uncorrelated sources causes an increase in overall level as a result of sound *power* levels adding together, not pressure levels. For a breakdown of 'sound pressure level', 'sound power level', and 'sound intensity level', and a fully detailed description of these phenomena, see Howard and Angus 2006, Chapter One.

> cancellation effects, the spatial variation in the sum of uncorrelated sources is usually much less that that of correlated ones, as the result only depends on the amplitude of the sources.

Masking

Masking can occur when listeners hear two or more sound sources simultaneously. The tonal qualities of discrete sounds often become difficult to discern if the individual frequency components of the simultaneously sounding soundwaves overlap. The degree of masking this overlap induces depends on the frequency content of the soundwaves in question, and their respective amplitudes. Sounds with similar frequency content battle for space in a mix, and recordists can minimize the effect of masking through equalization and panning, as well as other signal processing techniques, or by interjecting psychoacoustic cues.

However, though a particular instrument may be masked at any given point, the listener's mind understands that the obscured instrument still has a function in the mix. Imagine looking at someone on the opposite side of train tracks as a train passes quickly by. The person on the other side is only visible as the space between train cars allows, yet the mind infers the individual's physical presence even when unseen. Masking behaves in a similar manner, for an obstructed sound can "pop" in, and out, of direct perception as its masking track, or tracks, recede in volume or rest for tacit intervals.

From a psychoacoustic perspective, the mind anticipates auditory information it *expects* to receive. Although often considered undesirable, recordists sometimes use masking to create aggregate tones, that is, they blend sounds together to produce a hybrid instrument, usually for textural effect. For example, the track "There There (The Boney

King Of Nowhere)" from Radiohead's *Hail To The Thief* (2003) applies this concept from approximately 0:35 to 2:55 on the rhythm guitars, toms, and resonant feedback in the right speaker (set well back in the mix). The resonant tone, the rhythm guitar, and the toms, all combine to produce an aggregate texture which, while clearly a hybrid tone, takes on a singular role in the production at large.

PART TWO: HEARING

A comprehensive understanding of the musical functions of signal processing requires knowledge of the physiological determinants of human hearing. This section considers those aspects of hearing relevant to the perception of music, and it briefly outlines the functions of the outer, middle, and inner ear, while also introducing the concept of "critical bands" which, according to Howard and Angus (2006: 65), "is the single most important psychoacoustic principle for an understanding of the perception of music ... in terms of pitch, loudness, and timbre."

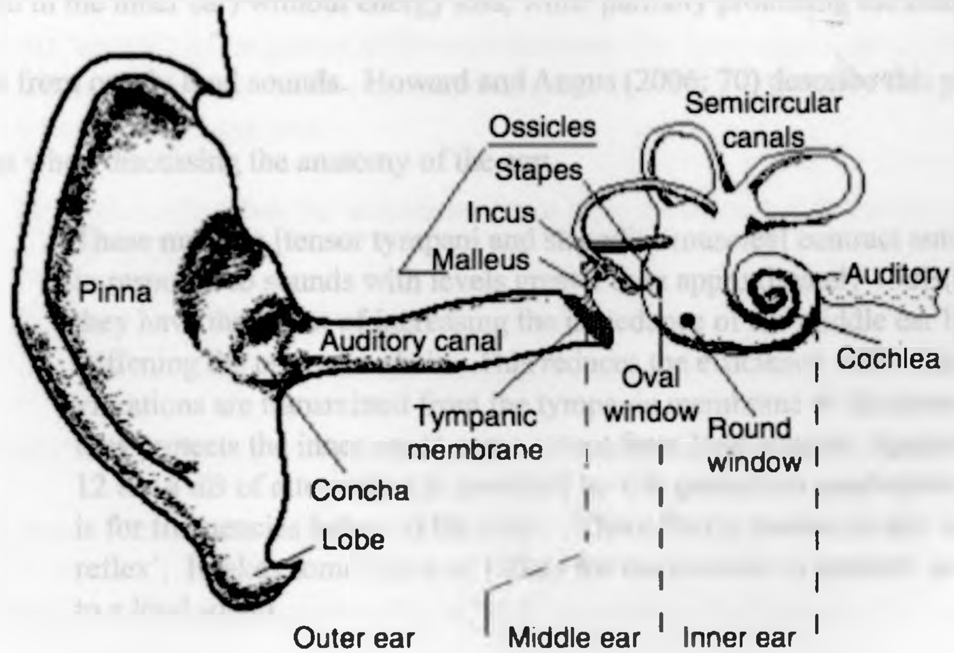


Figure 2.07. The main structures of the human ear showing the outer, middle, and inner ears.

The shape and anatomical composition of the outer ear (see Figure 2.07) determines how sound physically enters the ear canal. This acoustic effect serves two main purposes. It helps listeners localize sound sources, and it enhances certain frequencies. The outer ear modifies the frequency response of incoming sounds as the auditory canal has a natural resonance frequency of approximately 4 kHz. Sitting at the end of the auditory canal, the eardrum (tympanic membrane) converts acoustic pressure into mechanical vibrations in the middle ear.

The middle ear acts as a buffer between the outer ear and the sensitive inner ear. Its primary function is to transfer the motion of the eardrum to the fluid in the cochlea (located in the inner ear) without energy loss, while partially protecting the hearing system from overly loud sounds. Howard and Angus (2006: 70) describe this protection process when discussing the anatomy of the ear:

> These muscles [tensor tympani and stapedius muscles] contract automatically in response to sounds with levels greater than approximately 75dB(SPL) and they have the effect of increasing the impedance of the middle ear by stiffening the ossicular chain. This reduces the efficiency with which vibrations are transmitted from the tympanic membrane to the inner ear and thus protects the inner ear to some extent from loud sounds. Approximately 12 to 14 dB of attenuation is provided by this protection mechanism, but this is for frequencies below 1kHz only ... This effect is known as the 'acoustic reflex'. It takes some 60ms to 120ms for the muscles to contract in response to a loud sound.

The inner ear converts mechanical vibrations from the middle ear into nerve impulses, which the brain eventually processes as sound. The basilar membrane, also a part of the inner ear, acts as a frequency analyzer for input sounds, and the way in which this membrane functions determines the critical bandwidth at which subjective differences in perception occur.

Critical Bands

The basilar membrane separates the frequency content of incoming sound signals, and the "resolution" of human hearing refers to the point at which the basilar membrane can separate one frequency component from another. Three general ranges of tone differentiation exist:

> (i) "indistinguishable": when tones are close enough in frequency content (within approximately 12.5 to 15 Hz) that no differentiation in perception occurs;
>
> (ii) "rough": as frequency difference increases the fused tone changes to two separate tones but the line distinguishing the two is "rough" (they are not clearly isolated - beating); and
>
> (iii) "smooth": when the sensation of two separate tones persists (Howard and Angus 2006: 78).

Once tones reach the "smooth" state, they have achieved so-called "critical bandwidth" according to psychoacousticians. Of course, the specific point at which critical bandwidth obtains differs from one individual to the next, within a limited band. Every listener will experience an abrupt change in subjective response based on their own particular physiology, but the specific frequency and SPL at which this occurs varies slightly depending on personal physiology.

The importance of critical bandwidth should not be underestimated from a musical perspective. All complex musical sounds reduce to their individual frequency components, after all. As Howard and Angus (2006: 79) explain, "the resolution with which the hearing system can analyze the individual components or sine waves in a sound is important for understanding psychoacoustic discussions relating to ... how we perceive: melody, harmony, chords, tuning, intonation, musical dynamics, etc."

Pitch and Harmony

Musical instruments produce distinct pitches by creating regularly repeating acoustic pressure waves. Two components make up the nature of pitched and percussive waveforms, specifically, the time domain and the frequency domain. Time domain refers to the periodic or non-periodic nature of a waveform, whereas frequency domain refers to its spectral content (a *linear* spectrum contains harmonic components, while a *continuous* spectrum does not).

When a pitched instrument produces a particular pitch, it emits a range of frequencies all at once. This entire complex of frequencies is called a "composite" or "complex" waveform. The "fundamental" frequency of this composite waveform, which listeners subsequently register as pitch, corresponds with a letter name from the Western tonal system, given "well-tempered" instrumentation, which can also be described in Hertz (i.e.., "A above middle C," or, "A4 = 440 Hz").[11] In addition to the fundamental pitch, instruments produce a series of overtones, harmonic partials and formants, that relate to the fundamental frequency in particular ways.

---

[11] An exhaustive explanation of the harmonic series goes beyond the scope of this chapter. For a complete description of overtones/harmonic series please see The Cambridge History of Western Music Theory (ISBN 0 521 62371 5).

Defined as more than one pitch played synchronously, “harmony” is fundamental to tonal music. Psychoacoustic phenomena explain how listeners experience consonance and dissonance as crucial facets of harmonic perception:

> The development of western harmony follows a pattern where the intervals central to musical development have been gradually ascending the natural harmonic series. These changes have occurred partly as a function of increasing acceptance of intervals which are deemed to be musically ‘consonant’, or pleasing to listen to, as opposed to ‘dissonant’, or unpleasant to the listener. the psychoacoustic basis behind consonance and dissonance relates to the critical bandwidth, which provides a means for determining the degree of consonance (of dissonance) of musical intervals (Howard and Angus 2006: 139).

According to the Western tonal tradition, the “closer” two simultaneously occurring frequencies are to critical bandwidth, the more consonant they sound. Howard and Angus (2006: 141-142) offer an explanation as to how this distance applies to the ordering of interval perception based on a scale of consonance and dissonance:

> Musical intervals can be ordered by decreasing consonance on [a] psychoacoustic basis. To determine the degree of consonance of a musical interval consisting of two complex tones, each with all harmonics present, the frequencies up to the frequency of the seventh harmonic of the lower notes are found, then the critical bandwidth at each frequency mid-way between harmonics of each note that are closest in frequency is found to establish whether or not they are within 5% to 50% of a critical bandwidth and therefore adding a dissonance contribution to the overall perception when the two notes are played together. If the harmonic of the upper note is mid-way between harmonics of the lower note, the test is carried out with the higher frequency pair since the critical bandwidth will be larger ... The contribution to dissonance depends on where the musical interval occurs between adjacent harmonics in the natural harmonic series. The higher up the series it occurs, the greater the dissonant contribution made by harmonics of the two notes concerned ... (Howard and Angus 2006: 141-142).

By measuring the dissonance contribution of the harmonics in a two note chord, psychoacoustic theory explains the hierarchical ordering of dissonant intervals in western harmonic theory.

Timbre

Perhaps the most useful way to discuss timbre is by considering a pitch's attack, sustain, and release, that is, its so-called "envelope." The attack portion of an envelope determines the majority of its timbral characteristics, even though it usually lasts for only a few milliseconds. Psychoacoustic experiments demonstrate that listeners cannot reliably identify particular musical instruments when their attack and release stages are excised from the perceived pitch. Howard and Angus (2006: 227) describe the close relationship between timbral perception and a sound's attack envelope:

> ... if recordings of a note played on a violin open string and the same note played on a trumpet are modified to remove their onset [attack] and offset [release] phases in each case, it becomes very difficult to tell them apart ... the initial scraping of the bow on a stringed instrument, the consonant-like onset of a note played on a brass instrument, the breath noise of the flautist, the initial flapping of a reed, the percussive thud of a piano hammer and the final fall of the jacks of a harpsichord back onto the strings are all vital acoustic cues to the timbral identity of an instrument.

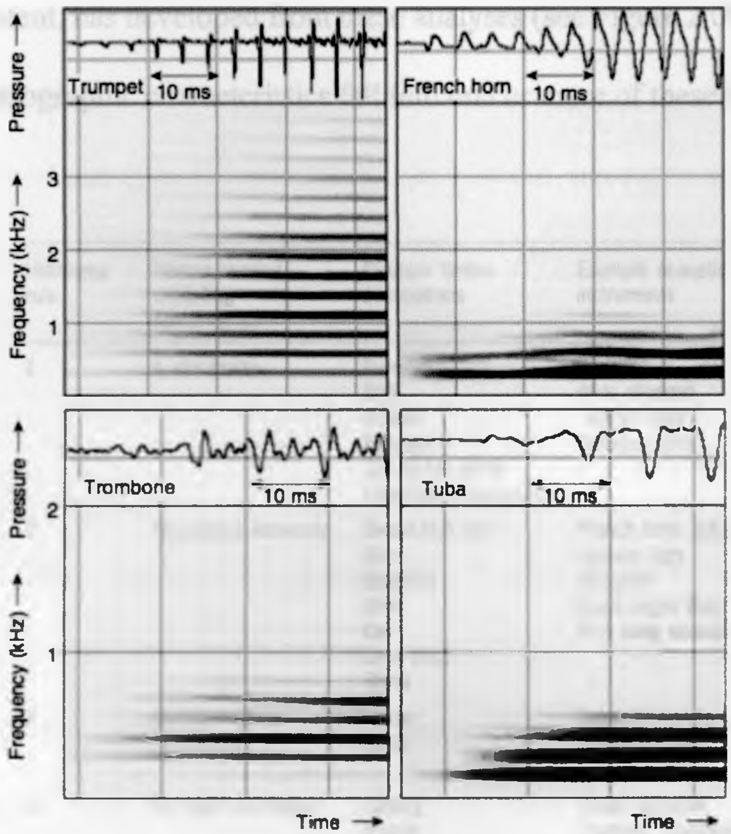


Figure 2.08. Waveform and spectrogram of note onset phase.

The notion of timbre can be fruitfully linked to the concept of critical bandwidth, in my opinion. As discussed earlier, the basilar membrane performs a frequency analysis on sounds entering the ear, and separates sound into its component frequencies (fundamental/harmonic partials/overtones/formants). Because the basilar membrane resolves harmonics below only the 5th to 7th partials, pscyhoacousticians hypothesize that lower harmonics play a distinct role in timbre perception, and they use spectrograms to model the process the human hearing system employs in order to visualize how the basilar membrane functions (see Figure 2.08). A summary of timbral categories, based on spectral content, has developed from these analyses (see Figure 2.09), and sounds with particular spectrographic characteristics fall into one or more of these categories.

Helmholtz rule	Human hearing modelling spectrogram	Example timbre descriptions	Example acoustic instruments
1	f_0 dominates	Pure Soft Simple Pleasant Dull at low pitch Free from roughness	Tuning fork Wide stopped organ flues Baroque flute
2	Harmonics dominate	Sweet and soft Rich Splendid Dark Dull Less shrill Bland	French horn, tuba Modern flute Recorder Open organ flues Soft sung sounds
3	Odd harmonics dominate	Hollow Nasal	Clarinet Narrow stopped organ flues
4	Striations dominate	Cutting Rough Bright Brilliant Shrill Brash	Oboe, bassoon Trumpet, trombone Loud sung sounds Bowed instruments Harmonium Organ reeds

Figure 2.09. Summary of the frequency domain properties as exemplified by the human hearing modelling spectrograms.

Of course, sound does not — cannot, in fact — exist in a vacuum. As sound moves through a medium, it also interacts with the physical properties of the space in which it moves. This has an obviously audible effect, and numerous conventions in the production process have emerged which all aim to emulate the physical behavior of sound in space, not the least of which being delay and modulation processing. As such, before finally explaining the techniques of delay and modulation processing, I shall have to briefly explain the natural basis for the sonic cues those techniques are used to embed within a recording, namely, room acoustics.

PART THREE: ROOM ACOUSTICS

The majority of music production and listening occurs in enclosed spaces. From recording on the studio floor to mixing in a control room, recordists must consider the acoustics of enclosed spaces when producing music. This section briefly examines the behaviour of sound in a room and how that behaviour affects the quality of *perceived* sound.

Direct Sound, Early Reflections, and Reverberant Sound

The soundwave from an instrument situated in an enclosed space reaches listeners' ears in several ways. First, it reaches the hearer directly, providing a signal uncontaminated by reflections, and a high degree of direct sound creates a clear and intelligible signal, which is particularly important for recognizing speech. Several milliseconds later, the first reflections of the direct sound arrive at the listener's position, and these sound waves are termed, aptly enough, "*early reflections*." Arriving later, and emitting from different directions than the direct sound, early reflections cue the listener to the size of the space in which a sound occurs, and to the location of the source within that space. Early reflections occur within the first 30-40 milliseconds of the initial occurrence of a sound. Reverberant sound, the sound that follows the early reflections, bounces around the room several times before arriving at the listener's position from all directions. This array of sound arrives at differing and overlapping times to create a dense set of delayed signals that combine to form the room's reverberation profile (see Figure 2.10).

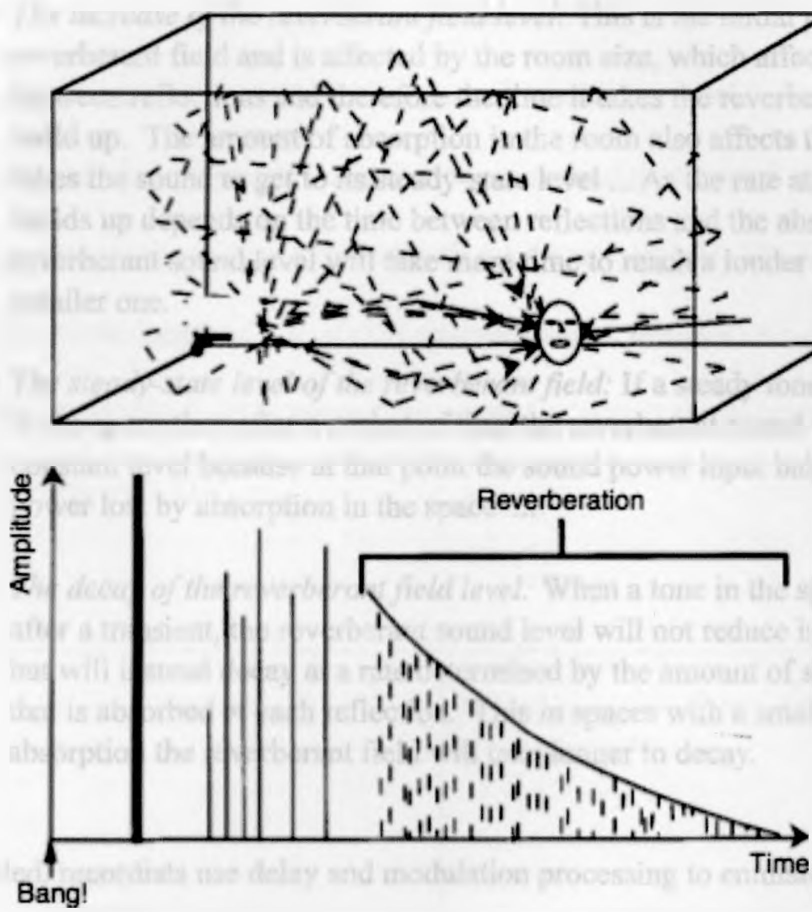


Figure 2.10. Reverberant sound in a room.

Reverberation Time

Reverberation time refers to the length of time it takes a sound bouncing around a room to go from onset to silence. Because sound loses energy each time it reflects off a surface, this energy loss causes the sound to die away gradually, and the size of a room, as well as the nature of its surfaces, determines how quickly the sound will decay. Smaller spaces have shorter reverberation times than larger ones, and Howard and Angus (2006: 269-270) explain the three aspects of the reverberant field that the size of the space determines:

- *The increase of the reverberant field level:* This is the initial portion of the reverberant field and is affected by the room size, which affects the time between reflections and therefore the time it takes the reverberant field to build up. The amount of absorption in the room also affects the time that it takes the sound to get to its steady-state level ... As the rate at which sound builds up depends on the time between reflections and the absorption, the reverberant sound level will take more time to reach a louder level than a smaller one.

- *The steady-state level of the reverberant field:* If a steady tone ... is played in the space then after a period of time the reverberant sound will reach a constant level because at that point the sound power input balances the power lost by absorption in the space ...

- *The decay of the reverberant field level:* When a tone in the space stops, or after a transient, the reverberant sound level will not reduce immediately but will instead decay at a rate determined by the amount of sound energy that is absorbed et each reflection. This in spaces with a small amount of absorption the reverberant field will take longer to decay.

As noted, recordists use delay and modulation processing to emulate the audible influence room acoustics exert over sound. Having surveyed the psychoacoustic bases for understanding that influence, I now turn my attention to explaining the technology recordists use to emulate it, namely, delay and modulation processing.

PART FOUR: DELAY, MODULATION/MODULATED DELAY, AND REVERB

All delay processors — including strict delay, modulated delay, and reverb emulators — use a delay line to suspend an input signal by a set amount of time, usually measured in milliseconds. This delay line takes an incoming audio signal and splits it so that one copy routes directly to the output of the unit (the "dry," or, "direct" signal), while the other copy is delayed by a buffer (i.e., digital memory, tape, etc.) for a specified amount of time before being sent to output.

Strict Delay

A delay line has at least three settings recordists control. 'Delay time' adjusts the number of milliseconds between the arrival of the direct and the delayed signal, while 'mix' regulates the amplitude ratio between both subsequent signals. 'Feedback', on the other hand, adjusts the length of time that the delayed signal remains active (that is, for repeated delays, it permits them to control the amount of output signal routed back to the input stage - see Figure 2.11).

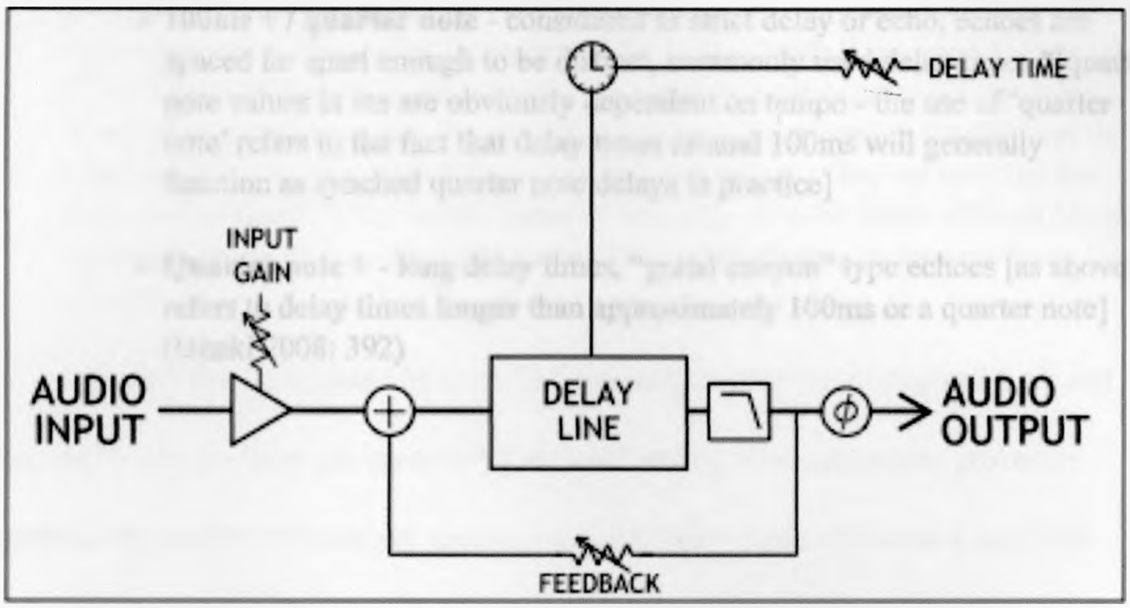


Figure 2.11. Typical signal flow of a simple delay line.

Psychoacoustically, humans have the ability to perceive minute variations in delay time settings. The difference of only a few milliseconds allows processed sound sources to function in a variety of ways in the context of a mix. Five main categories exist:

- **0-40 milliseconds (ms)** - very short, produces comb-filtering and alters timbral characteristics of instruments[12]
- **40-60 ms** - perceived as doubling
- **60-100 ms** - delay times over 60ms are perceived as distinct echoes (e.g. slapback)[13]
- **100ms + / quarter note** - considered as strict delay or echo, echoes are spaced far apart enough to be distinct, commonly used delay times *[quarter note values in ms are obviously dependent on tempo - the use of 'quarter note' refers to the fact that delay times around 100ms will generally function as synched quarter note delays in practice]
- **Quarter note +** - long delay times, "grand canyon" type echoes [as above, refers to delay times longer than approximately 100ms or a quarter note] (Izhaki 2008: 392)

[12] The interplay between the original and delayed signal causes constructive and destructive and destructive interference. The frequency response of a comb-filter consists of a series of regularly spaced spikes and dips, giving the appearance of a comb (see Figure 2.04).

[13] **Slapback** - A staple of 1950s rock is sometimes part of a contemporary mix: slapback echo ... Add a single audible echo between approximately 80 ms and 200 ms, and every note shimmers and pulses a bit, courtesy of the single, quick echo (Case 2007: 223).

Modulated Delay

Modulated delay lines form the basis of effects such as chorusing, flanging, and vibrato.[14] This kind of processing employs a low frequency oscillator (or, "LFO") to dynamically alter (or, "modulate") the delay time setting on a delay line. LFOs produce infrasonic frequencies below 20Hz which human ears interpret as pulsating rhythms rather than as discrete pitches. Alexander Case (2007: 213-214) explains the basic relationship between modulation and LFOs:

> The modulation section of a delay unit relies on a simple LFO. Instead of modulating the amplitude of a signal, as might be done in an AM (amplitude modulation) synthesizer, this LFO modulates the delay time parameter within the signal processor. Rate is the frequency of the LFO. Depth is the amplitude of the LFO. Shape, of course, is the type of LFO signal (e.g. sine, sawtooth, triangle) ... These three parameters give the recording engineer much needed control over the delay, enabling them to play it like a musical instrument. They set how fast the delay moves (rate). They set the limits on the range of delay times allowed (depth), and they determine how the delay moves from its shortest to its longest time (shape).

A delay line, modulated by a 10 Hz frequency, thus cycles through its peak and base amplitude ten times per second.[15] The "rate" setting on a modulation processor specifies the number of times per second a specific delay value cycles back and forth

---

[14] Phasing, usually identified as a modulation effect, must be considered a "special case" as a phaser uses an all-pass filter to delay wavelength to create the effect, though an LFO is applied to oscillate the phase shift, emulating modulation of the delay line. Phasing will be covered in greater detail in the next chapter.

[15] **Hertz** - A pure-tone sine wave consists of a simple, never-changing pattern of oscillation. Measure the length of time associated with each cycle to determine the waveform's period. Count the number of times it cycles each second for a determination of its frequency. Period is the time it takes for exactly one cycle to occur, with units of dimensionless cycles per second. Therefore, units for frequency live entirely in the denominator (per second, or /s) and have been given the alternative unit of hertz (Hz) (Case 2007: 8).

within the range set by the 'depth' control, and it is usually expressed in Hertz (referring to the Hertz of the frequency the LFO oscillates).[16] For example, a modulation effect with the delay time set to 10ms, the depth set at 50%, and the rate set at 10Hz will cycle the delay time setting on a delay line back-and-forth between 5ms and 15ms ten times per second (see Figure 2.12).

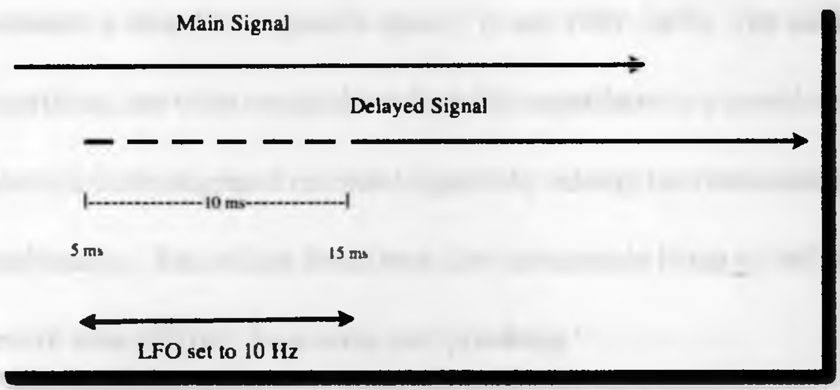


Figure 2.12. A modulated delay line with a delay time of 10ms, a depth of 50%, and a rate of 10 Hz.

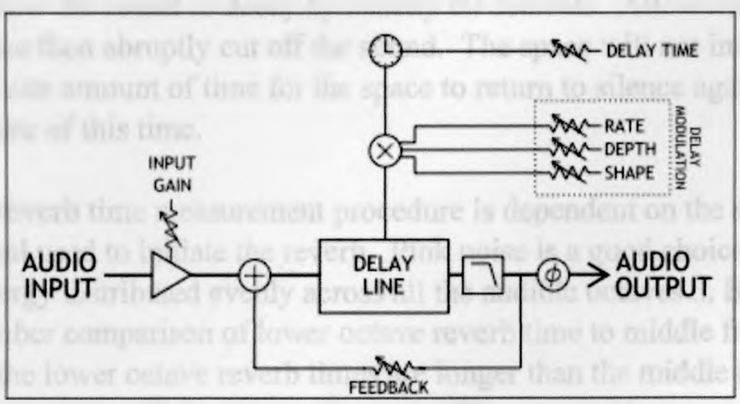


Figure 2.13. Typical signal flow for a modulated delay line.

[16] The 'depth' control of a modulated delay works as a percentage of the initial delay time setting (e.g. 10ms initial delay time setting with depth at 50% results in a delay line that cycles between 5-15ms) [10ms x 50% = 5ms].

Reverb

A full explanation of reverb would require a thesis all its own. The following brief description addresses only the basic nature, and the most common characteristics, of reverb, and reviews some of the more common considerations recordists entertain when they apply reverb in a mix.

Reverb is hundreds of delays blended together to create "the ambience cues that perceptually connect a sound to a specific space." (Case 2007: 269). The dimensions of a room, and its surfaces, are what create the delays that contribute to a room's reverberant qualities, and reverb units augment recorded signals by adding the resonance of physical space to a sound source. Recordists focus on a few measurable items to define the quality of a reverb: reverb time (RT60), bass ratio, and pre-delay.[17]

---

[17] **RT60** - The perceived liveness of an acoustic space is measured objectively by *reverb time, RT60*. Perhaps the most noticeable quality of a room's acoustics, reverb time describes the duration of the reverberant wash of energy. More specifically, it is the length of time it takes the sound to decay by exactly 60 decibels. Allow a sound to play in an enclosed space then abruptly cut off the sound. The space will not instantly fall silent - it takes a finite amount of time for the space to return to silence again. RT60 is the standard measure of this time.

Bass Ratio - The reverb time measurement procedure is dependent on the spectral content of the signal used to initiate the reverb. Pink noise is a good choice because it contains sound energy distributed evenly across all the audible octaves ... Bass ratio offers a single number comparison of lower octave reverb time to middle frequency reverb times ... If the lower octave reverb times are longer than the middle frequency reverb times, the bass ratio will be greater than one. The perceived overall warmth and low-frequency richness of a space is very much influenced by its bass ratio ...

Pre-delay - A gap in time exists between the arrival of the direct sound straight from the sound source to the listener and the arrival of any sound reflections of the reverberant wash of energy that follows. Pre-delay is the difference in time of arrival between the direct sound and the subsequent first associated reflection (Case 2007: 265-267).

SUMMARY

In this chapter, I have examined the psychoacoustic assumptions on which my model — and, as I hope is now clear, delay and modulation processing in general — is based. Recordists cannot physically situate one sound before another in a mix; they can only provide psychoacoustic cues which listeners physiologically interpret as spatial influence on a sound. Much as art historians are clear on the tools and techniques artists exploit to create three-dimensional perspective given only the limits of a two-dimensional canvas, so should analysts of Recording Practice and recorded musical communications understand the tools and techniques that recordists devise and exploit to establish the three-dimensional perspective of a mix. Situating tracks along the proximity plane of a mix is crucial in the creation of recorded musical communications, and delay and modulation processors provide recordists with the tools for invoking the psychoacoustic cues necessary for doing so. In the next chapter of this thesis, I will examine some of the conventional ways recordists use these tools for that very purpose.

CHAPTER THREE

PSYCHOACOUSTICS AND RECORDING PRACTICE

In Chapter Two, I examined a number of psychoacoustic cues which recordists use to create the proximity plane of a mix, and to establish motion along it. Much as cinematographers and film editors have established a number of common-practice techniques for establishing perspective and depth in film, so, too, have recordists created an entire lexicon of musical terms — an entire *praxis* aimed at processing sounds so they include psychoacoustic cues concerning distance and relative location — devoted entirely to creating perspective and depth on record. This chapter surveys the most common terms in this lexicon, and explains how they are achieved. Readers require this information to apply the model I present in this thesis, and to fully grasp its application in Chapter Four.

In this chapter, I remain chiefly concerned with delineating the broad contours of a particular musical practice, namely, spatialization at mix level. At first blush, this musical procedure often strikes novice recordists and analysts as too straightforward to qualify as anything so elevated as "musical practice." All but the most technically sophisticated research indeed implies that recordists simply position tracks along a proximity plane at some early stage in the production process, and then move on to the "true arts" of conveying idealized "live" performances to listeners vis-a-vis the medium of sound reproduction.

Most recordists describe depth as "the holy grail" of Recording Practice, as it were — something that only the most capable recordists convincingly achieve on a regular basis. Depth is, in fact, a difficult illusion to achieve for recordists of any stripe, even if they must do so each time they make a recorded musical communication. It is my simple hope that, by the conclusion of this chapter, readers will: (i) understand the complexity, and difficulty, of this task; and (ii) understand the complex procedures and techniques recordists have devised to create the psychoacoustic cues it requires (see Table 3.01).

Physical Dimensions	*Perceived Parameters*	*Artistic/Aesthetic Elements*
(Acoustic State)	(Psychoacoustic Conception)	(Resources for Artistic Expression)
Frequency	Pitch	Pitch Levels and Relationships—melodic lines, chords, register, range, tonal organization, pitch areas, vibrato
Amplitude	Loudness	Dynamic Levels and Relationships—program dynamic contour, accents, tremolo, musical balance, RDL
Time	Duration (time perception)	Rhythmic Patterns and Rates of Activities—tempo, time, patterns of durations
Timbre (comprised of physical components: dynamic envelope, spectrum and spectral envelope)	Timbre (perceived as overall quality)	Sound Sources and Sound Quality—timbral balance, arranging, performance intensity, performance techniques
Space (comprised of physical components created by the interaction of the sound source and the environment, and their relationship to a microphone)	Space (perception of the sound source as it interacts with the environment, and perception of the physical relationship of the sound source and the listener)	Spatial Properties—stereo location, surround location, phantom images, moving sources, distance location, sound-stage dimensions, imaging, environmental characteristics, perceived performance environment, space within space

Table 3.01. A table of the relationships between the physical, perceived, and aesthetic elements of sound.

As is evident in Table 3.01 above, psychoacoustic cues comprise the technical basis of all musical techniques in Recording Practice, not just those that situate, and resituate, sounds along the proximity plane of a mix. Though little definitive is written on this singularly modern musical practice, recordists have been at least intuitively aware of the role psychoacoustics play in allowing them to engage with the so-called "depth" dimension of a mix for the better part of a century now (see Sterne (2003) for more on historical approaches to the musical problem of "depth"). I agree with William Moylan (2007), in fact, that psychoacoustics should be understood to underwrite each aesthetic technique in the modern recordist's aesthetic toolbox, that is, that psychoacoustics comprises the technical basis of the recordist's musical discourse. What I particularly appreciate about this assertion is that it opens the possibility for *each stage* of the record-making process to play an influential, if not shaping, role in the construction of depth, as analysts conceive that musical practice.

In fact, all delay and modulation processing techniques — not just the ones examined in this chapter — occur in dialogue with the tracking and signal processing techniques which, in the first instance, create the tracks being mixed. To understand delay and modulation processing, then, I must situate them along a continuum of techniques which occur during every record production before mixing begins. Part One of this chapter, headed "Tracking & Signal Processing," thus surveys common ways that tracking and signal processing techniques, and technologies, "colour," or, distort, a sound source, that is, it examines the way recordists use certain tools and techniques to filter the psychoacoustic properties of a sound source so it best relates to a given multitrack production. Part Two, headed "Mixing," examines common methods recordists use to further refine the psychoacoustic cues captured during tracking and signal processing, in order to establish a proximity plane and emulate motion along it. In the following chapter, I will apply the concepts I examine in this chapter to a single track, establishing the viability of what I survey here as music-analytic categories.

PART ONE: TRACKING AND SIGNAL PROCESSING

The so-called "sound quality" of tracks is a product of the combination of several distinct variables. These *combined* variables include:

> (i) the timbral characteristics of the sound source itself, i.e.., the spectral profile of the sound source to be tracked;
>
> (ii) the acoustics of the tracking environment, i.e.., the spectral influence of room boundaries, dimensions and furnishings, on the spectral profile of the track to be tracked;
>
> (iii) the method of transduction, including direct-injection methods, microphone selection and placement, et cetera;
>
> (iv) the signal chain, i.e., the spectral influence of preamplifiers, compressors, equalizers, delay; and, most importantly, yet most egregiously overlooked,
>
> (v) the quality of AD and DA conversion, that is, overall distortions induced through digital sampling and quantization.[18]

Each of these variables is a complex entity unto itself, and requires sustained analytic attention on its own terms. Though it is well beyond the scope of this thesis to do this, I shall now briefly examine each "variable," in turn, with an eye toward establishing their position in Recording Practice at large, and the kinds of sonic information they are capable of interjecting in a recorded musical communication, before I turn my attention to delay and modulation processing *per se*. Each of the variables I list

---

[18] AD(C): Analog-to-digital conversion, DA(C) Digital-to-analog conversion

above, and examine below, imparts its own set of psychoacoustic cues onto a sound source, even while it refines the sound of cues interjected during preceding stages. It is, in my opinion, crucial to understand these "interjections" given that they structure the ideation and decision making process in which recordists engage to determine whether delay and modulation processing is necessary, and, then, to what degree.

Sound Source & Environmental Placement

Sound source selection requires that recordists consider: (i) which instruments and musician to use; (ii) which performance to select as the "best" take for the purposes of the multitrack project; (iii) which amplifier to use; and, then, (iv) where to place the musician/ instrument and amplifier (if applicable) within the broader tracking environment.[19] As mentioned earlier, a source *interacts* with the space in which it sounds, taking on numerous characteristics given so-called "acoustic interference." For our purposes, the most important of these "interferences" have to do with reverberation and echo. Early reflections, for instance, always arrive within milliseconds of the direct sound, and given that microphones capture these reflections, and that these reflections are primary in situating tracks spatially, great care must be taken when selecting the placement of the source in the room. Experienced producers understand the importance of spatial placement, and often work to balance direct signals with both early reflections

[19] This refers to a 'typical' recording setup from the 1950s through to the present day and applies most specifically in a rock/pop recording context. Some digital productions require only transduction for monitoring and occur almost entirely within the confines of a computer - this type of production, while fairly common by today's standards, it is the exception when considered as part of popular music's recording history.

and the reverberant field which accrues in the wake of sound's atmospheric disturbance within a particular room. In some cases, what would normally be considered an odd placement might be ideal to achieve a specific sound quality. Such determinations are entirely at the discretion of recordists.

Microphone Selection

Appropriate source placement conditions microphone selection, and the producer/ engineer will select microphones that enhance or complement the source in a particular way. Albin Zak (2001: 119) describes the importance of microphone selection in the recording process:

> The first step in the signal path is the microphone. In many ways microphones are the technological soul of any recording project; the effectiveness of all other tools and techniques depends upon the quality of the image that the microphone is able to deliver ... The microphone serves as an alchemic doorway between performance and text. Sound enters the doorway as fleeting vibrations in air pressure and is instantly transformed - or, more precisely, transduced - into a corresponding electrical signal that can be printed on some form of recording medium. But the process originally changes the sound in some way. the coloring effects of the microphone itself, along with those of its placement, must be appropriately matched with both the source material and the stylistic expressive intentions of the project - a task requiring aesthetic judgement and technical expertise. Decisions about microphone choice ... depend on a feel for the affective character of the sound in question and a sense of its role in the track.

No "rules" exist for selecting particular microphones for specific tasks, but general guidelines have developed as starting points for typical recording setups. Accomplished recordists like Mike Stavrou (2003) insist on treating each project as a unique set of circumstances. Stavrou has developed a "hardness" rating scale for microphones and sources (Stavrou 2003: 41-42). He suggests ranking microphones in a particular collection (e.g. a studio's microphone closet) from "hard" (1) to "soft" (10), in order to help recordists select appropriate microphones for the sources they will record. Stavrou likewise rates sound sources on a "hardness" scale of 1 to 10. Though he provides no definitive quantification for the scale, Stavrou's ranking taxonomy clearly measures transient response vis-a-vis crest factor, or a microphone's ability to accurately measure distance between transients and sustained information in a sound source in real-time. In any event, Stavrou ultimately recommends pairing a "harder" source with a "softer" microphone, and *vice versa*. Again, his "hardness" evaluations are, of course, utterly subjective, but a microphone's frequency response, operations principle (i.e., dynamic, condenser, ribbon), directional response, transient response, all combine to create an unique "frequency response" profile which reshape all signal passing through.[20]

Microphone Placement

Microphone placement works in conjunction with source location and microphone selection. Once recordists establish the optimal room placement for an instrument, and which microphone will capture the source/room combination according to the peculiar needs of each project, they must decide where to place the microphone relative to the instrument, the distance of the microphone from the source, and the angle of the microphone diaphragm vis-a-vis the center of the sound source. Because the largest

---

[20] A detailed review of microphone types and their specifics goes beyond the scope of this thesis. For an in depth review of microphones and microphone technique see Lewis (2010).

microphone diaphragms are actually quite small (several inches), and since transducers of limited proportions can pick up only an impression of what occurs acoustically, veteran recordists understand that at best they can merely represent, and not re-create, a musical event. Zak (2001: 110) describes the importance of microphone placement:

> The subtle art of microphone placement is nearly as important a factor in the rendering of the sonic image as microphone design, and recordists continually refine and expand their technique through experimentation. The placement determines the degree and type of coloration and defines the relationship between sound source and room sound. Considerations include not only the microphone's orientation to the sound source - as defined by distance and angle - and the acoustic characteristics of the room, but also the microphone's sound gathering properties ... A microphone's polar pattern, frequency response characteristics, and placement, are all factors in its timbral effect.

Preamplifier Selection

All electronic components alter signals passing through them to a certain degree. However, some devices "colour" the sound source, that is, they filter the spectral contour of the input source, more forcefully than others. A microphone preamplifier shapes the timbre of audio information passed through it, based on the configuration of its electronic components, that is, based on "signal path" (i.e., whether the preamp has input transformers, output transformers, tubes, discrete circuitry, balanced or unbalanced connections, *et cetera*). As discussed earlier, microphones transduce acoustic energy into electrical energy, and microphone preamplifiers boost electrical signals to "line level," making them useful for recording.[21]

Similar to Stavrou's "hardness" scale, preamps are often rated on a spectrum that ranges from "coloured" to "transparent," and these tonal differences shape the overall sound of tracks. Highly "coloured" preamps are said to have "character," while those which interject little as they boost signal are called "transparent," because their mediation is spectrally "transparent" or unobtrusive. While what specifically constitutes "colour" in this case, and how much distortion must occur before "colouring" can be said to occur, is, of course, open to interpretation, so-called "coloured" preamps typically introduce small amounts of harmonic distortion (which many listeners find pleasing), emphasize different parts of the frequency spectrum, and/or alter the signal's transient response, to a degree unmatched by so-called "transparent" preamps.

[21] A line level describes a line's nominal signal level as a ratio, expressed in decibels, against a standard reference voltage. The nominal level and the reference voltage against which it is expressed depend on the line level being used. While the nominal levels themselves vary, only two reference voltages are common: decibel volts [dBV] for consumer applications, and decibels unloaded [dBu] for professional applications. The reference voltage for the decibel volt (0 dBV) is 1 VRMS, which is the voltage required to produce 1 milliwatt [mW] of power across a 1 kilo-ohm [kΩ] load.[1] The reference voltage for the decibel unloaded (0 dBu) is the voltage required to produce 1 mW of power across a 600 Ω load (approximately 0.7746 VRMS) The most common nominal level for consumer audio equipment is -10 dBV, and the most common nominal level for professional equipment is 4 dBu. By convention, nominal levels are always written with an explicit sign symbol. Thus 4 dBu is written as +4 dBu. Expressed in absolute terms, a signal at -10 dBV is equivalent to a sine wave signal with a peak amplitude of approximately 0.447 volts, or any general signal at 0.316 volts root mean square (VRMS). A signal at +4 dBu is equivalent to a sine wave signal with a peak amplitude of approximately 1.737 volts, or any general signal at approximately 1.228 VRMS. (Ballou 2005: 761).

Audible differences between preamps are most noticeable when comparing preamps of very low and high quality components, with and without transformers, and with very different impedances. High-quality "coloured" preamps tend to have output transformers with iron, nickel, or steel cores, and Class A circuitry, while preamps with "low quality" components tend to veil or distort (in a non-pleasing fashion) information in the original signal due to less regulating circuitry.[22] Expert recordists understand how to select, and combine, microphones and preamplifiers to achieve specific sonic goals.

Signal Processing Chain (tracking stage)

In most instances, extensive signal processing takes place at the mixing stage. However, many recordists will track with compression and/or equalization on a majority of sources. When tracking, electric guitar "effects" like distortion, reverb, delay and modulated delay, are often printed, although, recordists do commonly split incoming audio to allow the capture of a "wet" and a "dry" signal, which allows them to blend, re-amp, or change the character of the "wet" signal at the mixing stage, when the broader

---

[22] Circuitry classes refer to the biasing of the transistor(s) or tube(s) during quiescent (no signal) conditions. In class A, there is one output device. It is always on, or conducting. During quiescent conditions, it is biased so that the output terminal (plate or collector) is at about 50% of the supply voltage. This offset voltage must be removed with a transformer or coupling capacitor so the output of the stage is 0V with no signal applied. As the input signal varies, the device conducts harder, toward saturation, or less, toward cutoff. As long as the stage is not driven into clipping, class A amplifiers can have very low total distortion, and NO crossover distortion, since they never turn off. Many audiophiles consider class A to be the best possible arrangement for audio reproduction (Ballou 2005: 590). A full discussion of an electronic component's sound altering qualities is well beyond the scope of this thesis but warrants further research, testing, and discussion.

spectral needs of the production at large are more fully evolved.[23] Recordists may include any type of signal processing in the tracking chain, but more often than not they minimize the amount of processing at the tracking stage (that is, with the exception of compression and/or equalization).

Medium: ADC/DAC and Tape

Recordists decide which medium will benefit a particular song at the outset of a project. However, the current technical circumstance of Recording Practice dictates that even records produced exclusively via analog equipment are usually consumed digitally. They thus go through an analog-to-digital conversion process regardless of inscription. Although all converters perform the same function, no two convertors sound, or transfer, alike. Converters impart a sonic signature on the material passing through them, in both analog-to-digital and digital-to-analog directions.[24]

Analog enthusiasts insist that tape simply sounds "better." While this statement is entirely subjective, it can be safely argued that recordings done to tape sound "different" from digital recordings in easily quantifiable ways, mostly having to do with spectral and

---

[23] Re-amping is the re-amplification of a 'dry' signal. Often used to re-amplify a 'dry' guitar take, a re-amping device receives line level signal from the recording device (e.g. hard-drive & converter, or tape machine & mixing console) and converts it into the kind of high-impedance signal a guitar amplifier requires. The signal from the converter is routed through the re-amping device, converted to a high-z signal, and is then fed into a guitar amplifier of choice. The resulting 'performance' emanating from the guitar amplifier's speakers is recorded again, thus allowing the possibility of multiple different-sounding takes of the same performance.

[24] The average listener may be unaware that the digital-to-analog conversion components in their consumer grade listening equipment affects the fidelity of their listening experience.

dynamic contour. When analog components begin to overload with signal, they create harmonic distortion (saturation), as opposed to the very non-musical harshness of digital clipping. A particular combination of tape machine, tape stock and tape speed (measured in *ips,* or, inches per second), brings a specific "sound" to the recordist's sonic palette. Tape adds a distinct equalization curve to the frequency content of an original signal, along with tape saturation (compression and harmonic distortion) when pushed to overload capacity.

PART TWO: MIXING

During mixing, recordists blend (or, "mix") sounds committed "to tape" in ways that attempt to satisfy the aesthetic priorities and needs of various interested parties: artist (s), record labels (if applicable), listeners, and, of course, recordists themselves. Listener expectations are conditioned by established practice, and professional recordists make it their job to familiarize themselves with those expectations. Roey Izhaki (2008: 33) describes some of the considerations mix engineers might make as they apply their craft:

> Do individual elements constitute the mix, or does the mix consist of individual elements? Those who believe that individual elements constitute the mix might give more attention to how the individual elements sound, but those who think that the mix consists of individual elements care about how the sound of individual elements contribute to the overall mix. It is worth remembering that the mix - as a whole - is the final product. This is not to say that the sound of individual elements is not important, but the overall mix takes priority ... [For example], it is extremely common to apply a high-pass filter on a vocal in order to remove muddiness and increase its definition. This type of treatment, which is done to various degrees, can sometimes make the vocals sound utterly unnatural, especially when soloed. However, this unnatural sound often works extremely well in mix context ... [Mixing] even goes into the realm of psychoacoustics - our brain can separate one sound from a group of sounds. So for example, while equalizing a kick we can

> isolate it from the rest of the mix in our heads. However, we can just as well listen to the whole mix while equalizing a kick , and by doing so, improve the likelihood of the kick sounding better in mix context. This might seem a bit abstract and unnatural - while we manipulate something it is natural to want to clearly hear the effect. The temptation to focus on the manipulated element always exists, but there's a benefit to listening how the manipulation affects the mix as a whole.

As the term "mixing" implies, various elements are blended together to create a recorded performance when mix engineers apply their craft. At this point, "the whole" takes precedence over its individual "parts," namely, tracks. Recordists place the separate strands created during tracking in an overall aesthetic context, and, when one identifies the various components that comprise the mix, and understands the role each component contributes to the final mix, new lines of analytic discourse become available. Bobby Owsinski (1999), for instance, offers a helpful model for discussing commercial productions in an analytically meaningful manner. In his model, Owsinski separates a mix into six distinct components: (i) balance; (ii) panorama; (iii) frequency range; (iv) dimension/depth; (v) dynamics processing; and, finally, (vi) interest.

Owsinski's Model: Balance

Balance usually refers to the perceived volume levels between the musical elements in a mix, but it can also represent the process of adjusting a song's arrangement to suit a particular mix. For example, when two or more instruments occupy similar frequency bands, they "fight" not only for space in the mix but also for the listener's attention. A well-written arrangement avoids this pitfall, but as recordists often have to deal with less-than-ideal tracks, they edit and re-arrange tracks for the benefit of the mix.

According to Owsinski, five main elements make up "the balance" of a rock record: (i) foundation; (ii) pads; (iii) rhythm (which I prefer to call counter-rhythm); (iv) lead; and (v) fills. "Foundation" generally refers to the bass and drums, but can also include rhythm instruments. "Pads," that is, sustained notes and chords, add texture to a mix. The "rhythm" element is any instrument that plays counter to the foundation element, and therethrough adds motion, and excitement, to the mix. The "lead" usually consists of a vocal line or a solo instrument, generally providing a song's melody. Finally, "fills" generally occupy spaces between lead lines (see Figure 3.01).

"Night Moves" Bob Seger	"Thank U" Alanis Morissette	"Two Pina Coladas" Garth Brooks
Foundation - bass, drums, acoustic guitar **Pad** - hammond organ **(Counter) Rhythm** - piano **Lead** - lead vocal **Fills** - background vocal answers and sometimes piano	**Foundation** - bass, drums **Pad** - synth in intro and chorus behind piano; different synths in chorus **(Counter) Rhythm** - piano; 'breath' sample in verse **Lead** - lead vocal **Fills** - guitar in 2nd verse	**Foundation** - bass, drums **Pad** - steel guitar **(Counter) Rhythm** - acoustic guitar and shaker **Lead** - lead vocal **Fills** - electric and acoustic lead guitar; steel guitar

Figure 3.01. Examples of arrangement elements in several popular recordings.

Owsinski's Model: Panorama

Panorama, sometimes referred to as width, denotes the placement of sonic elements in the stereo sound field. Panning instruments can create a sense of movement in a track, add clarity to an instrument by processing it to accommodate for the spectral and dynamic profiles of other tracks, and to expand the width and/or depth of a mix. The stereo panorama reflects the horizontal plane of the imaginary soundstage and, based on conventional practice, listeners expect the most important instruments to occupy the center channel of a mix. While there are no "rules" in mixing, convention states that lead tracks (vocal or otherwise), kick drum, snare drum, and bass sit at (or very near) the center position. Hodgson (2010: 165) discusses panning conventions in modern Recording Practice:

> Though there can be a wide degree of variance between where, exactly, the kick drum, snare drum and bass tracks are anchored in a mix, they almost always remain within a few degrees of center. When these tracks are not centered, mixers are aware that listeners expect them to be and, so, they situate them elsewhere in a mix to achieve some psychoacoustic or aesthetic effect. Lead vocal tracks, on the other hand, are almost without exception panned to the front-and-center.

Owsinski's Model: Frequency Range

Recordists must consider the frequency content of a mix in its entirety, as well as the frequency content of each individual track within it. As previously discussed, mixers balance individual instruments based on how they affect the whole mix, and every recorded instrument has particular frequency bands that will either heighten, or diminish, its prominence in the mix. For example, a bass drum's lowest, most powerful frequencies will be found between approximately 80 Hz to 100 Hz, its hollowness in the vicinity of 400 Hz, and the sound of the beater hitting the skin somewhere between 3 kHz to 5 kHz. To "clarify" the kick, then, recordists often boost its amplitude in the "attack" range, just as they may boost its lowest octave, namely, 80 to 100Hz, to increase the force of each hit.

A mix's entire frequency spectrum has six distinct areas, each one contributing differently to the mix. The sub-bass area contains very low frequency information, approximately 16 Hz to 60 Hz, and gives the mix a sense of power as these frequencies are "felt rather than heard." The bass ambit of frequencies operate in the 60 Hz to 250 Hz range, and contain the fundamental notes of the rhythm section. The low-mid frequencies, ranging from 250 Hz to 2 kHz, contain the low order harmonics of most instruments. High-mids, approximately 2 kHz to 4 kHz, contain the frequencies most associated with speech recognition and, as such, are an area to which humans are particularly sensitive. Spectral content falling between 4 kHz and 6 kHz affects proximity perception and is often referred to as the 'presence' band. Occupying the uppermost bands of the spectrum, 6 kHz and above, lie the frequencies that help add to a sound's clarity which Owsinski refers to as having the "brilliance" or "air" of tracks.

Owsinski's Model: Dimension/Depth

Dimension alludes to the ambient field in which a mix, or a particular track, sits. The physical space captured at the tracking stage contributes to this ambient field but creating and enhancing depth through the use of reverbs, delays, and modulated delays, is more common at the mixing stage. A recordist will add dimension to a track in order to create the illusion of space, add excitement, make a track sound bigger/wider/deeper, or to give the impression that a particular instrument is farther away, that is, to move it back along the proximity plane of the mix. Dimension relates most directly to reverbs and delays, a subject I examine in greater detail below.

Owsinski's Model: Dynamics Processing

Dynamics processing refers to all manners of compression, limiting, expansion, and gating. Compression, the most common of all signal processing found in popular music production, deserves a full study on its own, but for the purposes of this thesis, I will provide a brief description to make readers aware of its ubiquity and general characteristics. Compression furnishes the recordist with an automated level control, and Owsinski (1999: 48) discusses how a compressor operates:

> Compression is an automated level control, using the input signal itself to determine the output level. This is set by using the threshold and ratio controls. Compressors work on the principle of gain ratio, which is measured on the basis of input level to output level ... For example, this means that for every 4dB that goes into the compressor, 1dB will come out for a ratio of four to one (normally written as 4:1). If a gain ratio of 8:1 was set, then for every 8dB that goes into the unit, only 1dB will come out. Although this could apply to the entire signal regardless of level, a compressor is usually not set up that way. A threshold control determines at what signal level the compressor will begin to operate. Therefore, threshold and ratio are inter-related and one will affect the way the other works ... Most compressors also have attack and release parameters. These controls determine how fast or slow the compressor reacts to the the beginning and end of the signal ... When a compressor operates it actually decreases the gain of the signal so there is another control called make-up gain or output, which allows the signal to be boosted back up to its original level or beyond.

A compressor, then, gives the recordist control over a sound's dynamic envelope and perceived level. By reducing the dynamic peaks of a track, its overall average amplitude can be (seemingly paradoxically) raised, which often has the effect of increasing the sound of the space the instrument was recorded in since it results in amplification of low level information in the input signal.

Owsinski's Model: Interest

Ideally, a mix captures and holds a listener's attention from beginning to end, but in order for a mix to fulfill this ideal, it must engage the listener at an emotional level and withstand repeated listening. Recordists often manage this feat by identifying which instrument is most important to the 'sound' of the recording so they can emphasize it in the mix. That is, they build the mix in a way that supports and emphasizes the audible character of each identified instrument. It is simply beyond the scope of is thesis to engage this mix element in greater detail.

Owsinski's Model: Summary

Recordists make aesthetic decisions throughout the record making process that culminate in a product primed for public consumption. They base their decisions on personal experience and aesthetic convention. Part One of this chapter has broadly identified the decisions recordists make when assembling tracks for mixing, and the possible aesthetic consequences which can thereby accrue. I now turn my attention to providing a similar survey for the mixing process in Part Two below.

Mixing Continued: Delay, Modulation, and Reverb Processing During Mixing

Delay and modulation processing are the most referentially fixed of any processing techniques, likely due to the fact that they are also, generally, the most obvious (Hodgson 2010). In fact, delay and modulation often serve as markers of genre. From slapback echo on rockabilly records from the 1950s, to multi-tap echoes subdividing the beat in modern dance and electronic music, reverb, delay, and modulation processing remain ubiquitous techniques on modern Recording Practice. The balance of this chapter examines conventional applications of delay, modulation, and reverb processing techniques during mixing in a roughly modern rock context. Chapter Four will concretize the musical concepts I raise in Part Two of this Chapter through a "close reading" of delay and modulation processing on The Police's "Bring On The Night" (1979).

Slapback

Indicative of rockabilly and the emerging rock 'n' roll sound of the 1950s, slapback refers to a single audible echo that arrives approximately 80ms to 200ms after the original sound. Elvis Presley's "Sun Sides" records prominently feature slapback echo which, in turn, prompted John Lennon to use it conspicuously on many of his solo recordings. The ubiquity of slapback echo on 1950s recordings coming out of Sam Phillips' Sun Records studio root this particular processing technique in that time and place. Alexander Case (2007: 224) explains how slapback was achieved prior to digital signal processing:

> Before the days of digital audio, a common approach to creating this sort of effect was to use a spare analog tape machine as a generator of delay. During mixdown, the machine is constantly rolling, in record. The signal is sent from the console to the input of the tape machine ... using an echo send or spare track bus. That signal is recorded at the tape machine and milliseconds later it is played back. That is, though the tape machine is recording, it remains in *repro* mode so that the output of the tape machine is what it sees at the playback head ... the signal goes in, gets printed onto tape, the tape makes it's way from the record head to the playback head (taking time to do so), and finally the signal is played back off tape and returned to the console. The

> result is tape delay. The signal is delayed by the amount of time it takes the tape to travel from the record head to the repro head. The actual delay time then is a function of the speed of the tape [in ips] and the particular model of tape machine in use (which determines the physical distance between the two heads.

Analog tape delays are still in use, and when recordists choose analog tape as a delay source they do so for aesthetic reasons. As previously mentioned, each combination of tape, machine, and speed has its own sound. Tape delays often have variances in pitch ('warble' and 'flutter') which accrue given imprecise and irregular tape speeds, along with the low-frequency bias of tape which causes the delay line to progressively filter more high-frequency content from each subsequent echo (Hodgson 2010: 126). Slapback connotes either 1950s rockabilly and rock 'n' roll, or it functions as spatial nuance on an otherwise static track.

ADT (Automatic/Artificial Double Tracking)

Recordists often double-track lead vocals in order to thicken their texture. Prior to ADT, this process was a tedious and manual one. ADT is first found on The Beatles 1966 album *Revolver*. The process was invented by EMI maintenance engineer Ken Townshend when he realized that copying an existing track to a tape machine equipped with varispeed would allow the duplicated track to be slightly sped up and slowed during playback, thus mimicking the imprecision of a manually doubled second take. George Martin and Geoff Emerick emulated Lennon and McCartney's manual doubling by panning the input and doubled signals to opposite sides of the stereo image. They discovered an additional aesthetic effect when they panned the doubled track to the same horizontal position in the mix as the input signal to achieve an effected form of ADT. Ryan and Kehew (2008: 297) explain this phenomenon and how its use was extended to instrumental parts:

> When the original signal and delayed signal were panned to the same spot in the stereo picture, a distinctive sound emerged. The presence of two separate sounds could be discerned but the signal sounded affected: it did not sound entirely like natural double-tracking. This of course, was one of the qualities that greatly attracted The Beatles to the effect, aside from its time saving benefits. However, when the two signals were panned to different parts of the stereo image, the double-tracking effect could be quite convincing indeed. A listen to the brass in "Savoy Truffle" illustrates this nicely; with the original signal panned to one side and the delayed signal paned to the other, the illusion of two separately overdubbed parts was remarkable. The vocals on "Ob-La-Di-Ob-La-Da" and "Birthday" were handled similarly. Quite often on later Beatles records, this effect was used to create a lush stereo image, an especially useful effect when only four tracks were available.

ADT's delay time characteristics fall into the 20-50ms range and, as previously noted, its effectiveness as a convincing "second take" depends on its (and the input signal's) placement in the stereo image.

Multi-tap Delay/Echo (synched)

A multi-tap delay offers precise control over the number of echoes a delay line produces, their amplitude relative to the input signal and to one another, the panning location of each echo, their spacing as a function of time, and their frequency content. Exclusively digital, multi-tap delays first entered use in record production in the mid-1970s, but have more recently achieved distinction in modern electronica.[25] Jay Hodgson (2010: 138) explains the use of multi-tap echoes in contemporary productions:

> ... electronica productions feature multi-tap processing in a primarily rhythmic role now, with recordists multi-tapping irregular sequences of echoes to generate rhythmic propulsion and momentum for tracks. Though examples of multi-tapping abound on modern electronica records, Tosca's "Suzuki" provides an exceptionally clear demonstration. The track begins with a single harmonic, percussively plucked on an electric bass, followed by a sequence of regularly diminishing echoes. When the harmonic repeats, however, it is followed by a multi-tapped sequence of echoes that randomly ping-pong back-and-forth across the stereo spectrum: right-right-left-right-left-right-left-left-right-left-left-left-left, et cetera. Compared with the regularized echo-delay heard on, say, the electric bass line which introduces Porno for Pyros' "Pets," the multi-tapped echoes on "Suzuki" clearly belong to an entirely different genus.

Echo/Delay (unsynched)

In modern Recording Practice, delays of longer than 50ms tend to be synchronized with the tempo of tracks. However, recordists just as often require a non-synchronized delay. Most often, the unsynched delay line is used to emphasize a particular part, or instrument, by placing it in a distinct spatio-temporal continuum, as indicated by its unique echo and reverberation profile. Hodgson (2010:135-136) provides an example of unsychronized delay in a modern rock context:

> While the Edge's guitar playing in general demonstrates basically every delay processing technique in the modern rock recordist's toolbox, his guitar work for "Stuck In A Moment You Can't Get Out Of," ... provides a particularly clear demonstration of the unsynced principle at work in modern rock. Just before the song's pre chorus at 0:29, the Edge's dry guitar is suddenly

---

[25] The first multi-tap delays were rack mounted hardware units equipped with their own analog-to-digital/digital-to-analog converters which allowed their integration into standard tape/console (analog) based recording setups.

> delayed. The dry track remains on the left side of the stereo plane while the delay line is bussed (sent) to an open channel on the opposite side. Because the Edge subdivides the basic tempo of the track with a sequence of straight quarter- and eighth - note arpeggios at this point, producer Daniel Lanois had little choice but to unsync the delay line. Synchronized echoes would have overlapped the dry line, potentially inducing masking and, perhaps most egregiously in the soft rock world that U2 dominates, including inappropriate dissonances each time the harmony changes.

Flanging

Flanging, the result of modulating a delay line with a short enough delay time setting that comb filtering obtains (from roughly 1ms - 75ms), is often used by recordists to emphasize a particular instrument by altering its spatial location vis-a-vis the aural perspective which each mix construes (Hodgson: 2006). Due to the rate and depth of the modulation (as discussed in Chapter Two), the series of peaks and dips that make up the subsequently induced "comb filter" sweep up-and-down the frequency spectrum, thus creating the flanger's characteristic "whooshing" sound. Roey Izhaki (2008: 403) explains the flanging effect and its control parameters by way of analogy:

> ... the flanging effect is like a siren where the depth [parameter] determines how low and high the siren goes and the rate [parameter] determines how quickly it goes. The feedback can be thought of as a resonance controller, the higher the feedback the more the resonance.

Flanging receives its name from the metal flanges on a tape machine used to hold tape-reels in place. Before the advent of digital signal processing, recordists would apply pressure to the flange with their fingers to create the flanging effect. By varying the amount of pressure applied to the flange, they altered the delay time and induced the familiar 'whooshing' sound. Although found on records by Les Paul, and other pioneers, flanging entered widespread use on rock recordings in the late 1960s and 1970s (e.g. Jimi

Hendrix, "Axis, Bold As Love" [1967]), not only vis-a-vis particular instruments but also entire sections of a mix (e.g. Small Faces, "Itchycoo Park" [1967]).

Chorusing

Chorusing creates cyclic variances in pitch, or a sequence of pitch-modulations, by utilizing a series of modulated delay lines, and is generally agreed to "push" tracks back along the proximity plane. Through the use of several different delays, generally set in the 20ms to 50ms range, chorus emulates exactly what its name implies: a unison chorus of instruments (including voice). The original signal shunts directly to output, along with the modulated delay lines the chorus process creates, resulting in a chorus' characteristic "warbling" effect. Chorused electric guitars began to appear *en masse* in the Top 40 soundscape in the late 1970s and early 1980s, "which many historians, and critics, consider to have been a golden age for the chorus effect" (Hodgson 2010: 143).

From a mixing perspective, chorusing can create the impression of multiple performances, each with a slightly different pitch and timing, which recordists often apply to backing vocals and string parts. Chorusing also frequently appears on synthesized organ and string parts, to ensure they do not encroach on the proximal location of front-and-center mix elements, and on acoustic guitar, bass guitar and, sometimes, vocal tracks, to compensate for slightly out-of-tune deliveries.

Phasing

Generally considered part of the modulation processing genus, phasing actually employs a series of all-pass filters, as opposed to delay lines, to reshape the frequency content of an input signal. An all-pass filter shifts the phase of the input signal to delay that signal at a rate determined by wavelength. Hodgson (2010: 144) explains how all-pass filters react to different frequencies:

> Low frequencies, which have longer wavelengths, thus shift (delay) at a slower rate than do higher frequencies ... Notches, that is, muted frequency ranges, subsequently accrue at frequencies where the phase-shift induced by the all-pass filters is precisely 180-degrees. This emulates the effect of comb-filtering and, when a modulating LFO is applied to the phase rate of those filters, the notches subsequently sweep back-and-forth across the frequency spectrum.

Phasing, often mistaken for flanging, is generally a subtler effect. Unlike a flanger, the peaks and dips produced by a phaser's comb-filter do not occur in harmonic series. Often used to process electric guitars, phasing creates cyclic variance on a track, generating subtle spatial front-to-back motion along the proximity plane. Like other delay and modulation effects, setting the phaser's rate to coincide with, or subdivide, the tempo of the song is standard practice, though equally often recordist's opt for slight tempo misalignment.

Tremolo

Tremolo, a form of amplitude modulation, cycles an input signal's amplitude between a low and high setting determined, again, by "rate" and "depth" controls on a connected LFO. Rate, expressed in Hertz, governs the speed of the modulation and depth, conveyed as a percentage, determines the intensity of attenuation. Like other modulation effects, tremolo can bring a sense of motion to an otherwise static track, but most often it induces a degree of spatial dynamism, pushing tracks slightly back from the front-and-centre of a mix, especially in relation to "dry" tracks. Commonly found on instrumental surf rock tracks from the 1960s, for instance, a pronounced tremolo effect can allude to this genre (particularly when combined with copious amounts of spring reverb).

Ducked Reverb

Ducking involves lowering the amplitude of one signal based on the output of another signal. To achieve a ducked reverb, a processing technique so common it defies genre characterization, recordists insert a gate (set to ducking mode) on the track requiring attenuation: in this case a reverb auxiliary bus (see Figure 3.02).[26] A reverb unit

---

[26] **Gate**: As compressors attenuate signal above the threshold, noise gates attenuate signal which registers below the threshold. Unlike compressors, however, these gates attenuate signal by a fixed amount, called the range. Recordists chiefly use gates to reduce the input signal to silence at quiet intervals, which can require a range of more than 80 decibels of immediate reduction, though the device has plenty of other established uses. Aside from 'attack' and 'release' settings, gates also typically feature a 'hold' setting, which determines the length of time - usually anywhere from zero to three seconds - the gate remains active once the signal which triggered it has subsided under the threshold (Hodgson 2010: 86-87).

reacts to a "send" from the input channel — a lead vocal track, for example — and can sometimes intrude on the clarity of the mix. To remedy this, recordists will duck the reverb output under the vocal take, that is, when the lead vocal track's amplitude exceeds the threshold of the gate it reduces the reverb channel's output, thus clearing space for the vocals. The dimensional, or spatial, effect attributed to reverb (i.e. its ability to move things 'back' in the mix and create space around the input signal) is lessened by the ducker: when the vocal track subsides under the gate's threshold, the reverb tail swells into the pauses between vocal phrases, creating a timbral modification to the "release" profile of a track's envelope rather than the entire ADSR profile. Ducked reverb often appears on snare drum tracks, as well as vocals, in a multitude of pop genres.

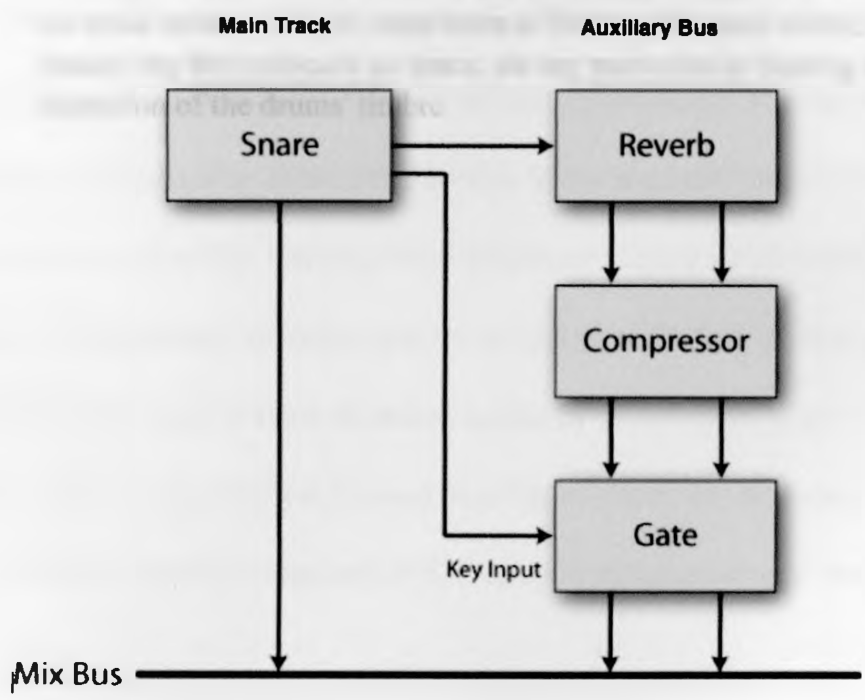


Figure 3.02. Signal routing for a ducked reverb on an auxiliary bus.

Gated Reverb

Gated reverb, famously used to achieve Phil Collins' massive drum sounds on "In The Air Tonight", allows a recordist to capture the initial burst of reverb created by an input signal and truncate the reverb tail. The result is achieved by inserting a gate in the signal chain after the reverb output and setting the gate threshold so it opens with the initial reverb burst yet closes on the tail.[27] Albin Zak (2001: 80-81) describes the effect's function in the mix as well as the listener's psychoacoustic perception of it:

> The effect in this case is one of textural clarity ... The gated ambience on the drums resolves a physical contradiction: it evokes the size associated with large, open ambient spaces while at the same time confining the sound to a clearly delineated place in the track. Limiting each eruption of sonic intensity to a short burst solves the problem of preserving textural clarity without diminishing the ambient drums' visceral power ... With the ambient decay continuously truncated ... the strange behaviour of the sound in this atmosphere dictates the terms of our acoustic perception. Because it makes no sense in terms of our experience of the natural sound world, rather than perceiving the ambience as space, we are reoriented to hearing it simply as an extension of the drums' timbre.

---

[27] The signal routing for a gated reverb looks identical to that of the ducked reverb (Figure 3.03) - the audible difference between the two effects depends on the gate's settings.

CHAPTER 4

APPLYING THE ANALYTIC MODEL: BRING ON THE NIGHT[28]

This chapter applies the analytic model I have thus far developed vis-a-vis The Police's "Bring On The Night." To be clear, I will focus explicitly on examining the musical function of delay, modulation and reverb processing throughout the track. In the the next chapter, I will consider various directions in which my model might be taken for future research, to ensure its broader scholarly generalizability.

INTRODUCTION (0:00-0:26)

Foundation: Hi-Hats & Guitar

The hi-hats introduce the production, establishing its meter and tempo through a straight sixteenth-note pattern with heavily accented downbeats. The higher amplitude of the accented downbeat hits creates considerable "room response" in the recorded signal which, in turn, disperses the increased high-frequency content which transfers to the delay line. On first blush, the delay processing applied to the hats produces a comb-filter effect similar to the kind of tonal distortion caused by phase-shifts induced during tracking. However, repeated and focused listening confirms that the sound is, in fact, a psychoacoustic composite comprised of (i) a delayed signal arriving in the right channel

---

[28] Album: Regatta De Blanc (1979), Key: G, Time Signature: 4/4, Tempo/BPM: 110

several milliseconds after (ii) the direct signal.[29] Along with the room information captured during tracking, the delay processing ultimately functions to situate the hi-hats slightly back, and to the right, along the horizontal and proximity planes of the mix, providing a "sonic anchor" for the broader mix at large.

The frequency content of an individual track will compete for space in the mix if similar frequencies exist in another track sounding at the same time. The effect is compounded if the tracks in question occupy the same horizontal position, that is, if they are co-located along the horizontal plane. Recordists counter masking using equalization, dynamics and spatial processing, to adjust a track's size and location in the vertical, horizontal, and proximity planes, and through differential placement in the arrangement (i.e.., ensuring conflicting spectral profiles do not sound at the same time). In this case, the hi-hat's spectral content is mainly comprised of frequencies between 200 Hz - 300 Hz, and 4 kHz - 14 kHz, with the majority of the energy in the latter band. No other instruments occupy the same upper bands, at the same times. While frequency range cross-over between tracks is inevitable, recordists mitigate unwanted masking effects by optimizing the frequencies that play a crucial role in creating the important timbral qualities of each respective instrument in the mix (hence, the composite spectral focus of the hi-hat track).

The majority of the guitar's frequency content lies between approximately 100 Hz - 260 Hz (where the "body" and "fullness" of the guitar's timbre is found) and between

---

[29] An interview with Andy Summers in an issue of Guitar Player magazine from September 1982 indicates that Stuart Copeland often employed the use of an Echoplex (tape delay) for aspects of his distinct drum sound.

approximately 900 Hz - 2 kHz (where the percussive "attack" and "presence" of the guitar's timbre are situated). While some spectral overlap between the hi-hats and the guitar accrues in the 200 Hz - 300 Hz range, it serves to reinforce the foundation of the mix as the majority of the masking occurs on the emphasized downbeats.

A subtle modulated delay line gives the guitar track motion along the proximity plane. The subsequent resonant sweep (or "whooshing" sound) of the process reveals the modulation to be flanging (rather than chorus or phase). The rate of the flanger cycles at approximately 0.25 Hz (110 BPM / 60 = 1.8 BPS / 8 = 0.225 Hz), the depth control is set at under 1ms, and the flanger's feedback is under 20%.[30] The result is a subtle flanging effect running the full cycle of its sweep over two bars. The flanging effect and the "room" sound situates the guitar within its own unique space. Discerning the order in which the flanger and reverb processor occur in the signal chain is difficult. Is the reverb flanged or does the flanging take place within the space created by the reverb? Or does the entire mix receive flanging treatment? Because of the flanger's subtlety, and the overlap of 'space' between each respective track, any of the three aforementioned scenarios seem possible, but without session documentation or access to the mix engineer (both of which are, of course, exceedingly rare), the listener/analyst is left to decode the signal path themselves. Experimentation and re-creation of the various signal chains offers the only way to test which scenario is most likely.

---

[30] These settings can be confirmed by generating a test tone (e.g. a sine wave at 1 kHz) and inserting a flanger with the above settings into the chain.

Counter-Rhythm: Kick Drum

The kick drum provides a counter rhythm to the quarter-note downbeat emphasis of the hi-hats and guitar when it appears on the upbeats of beats 1 and 3. Accenting upbeats is common practice in reggae, a Jamaican popular music style from which The Police admittedly borrowed often. In fact, the album title, *Regatta De Blanc,* is meant to signal their indebtedness to the pioneering work of Bob Marley, Peter Tosh, Toots and the Maytals, et cetera, roughly translating as "White Reggae."

Considerable reverb processing draws further attention to the kick drum. The wet/dry signal ratio across the stereo spectrum defines the sound and place of the kick drum. The dry kick signal sits slightly right of center in the mix, and the wet tail of the reverb pans from the dry signal across the stereo field to the extreme left. As the dry signal emerges, attention to the right channel reveals that what might initially be mistaken as reverb is, in fact, a combination of snare rattle and room ambience captured during tracking. The fact that only the reverb effect pans across the spectrum, and not the dry hit, indicates the use of an auxiliary bus for the reverb effect. The reverb's decay setting more or less coincides with the tempo of the song, as the tail lasts approximately the length of one half-note.

The psychoacoustic information produced by the reverb implies the size of the space in which the kick drum sits. The reverb's frequency and energy content diminishes over the length of the tail; the effect is psychoacoustic mimicry of the inverse square law.[31] Because the reverb process breathes in time with the tempo of the song, and because the process is created by the acoustic qualities of the drum itself, some listeners may perceive the reverb process as not reverberations but, rather, simply a component of the kick drum's broader timbre.

Two possible signal-processing paths exist for achieving the kind of reverb effect found on the kick drum. A mono, or stereo, bus panned left with a reverb insert set to 100% wet and a very short pre-delay setting, could achieve this effect. However, so, too, could a ducked reverb line. A ducked reverb line also employs the use of an auxiliary bus. In this case, though, the bus's inserts would contain the reverb processor followed by a ducking unit (compressor/gate/dedicated ducker). Instead of using a pre-delay setting on the reverb itself, which would not duck early reflections but, rather, simply delay their onset — some reverb units do not have this control parameter — the ducking unit is then keyed to the dry kick drum track.

---

[31] **Inverse Square Law**: A sound source radiating energy into three-dimensional space produces an intensity that falls off in inverse proportion to the square of the distance from the observer to the source. This means a reduction of sound pressure of 6 decibels for each doubling of distance if the source is in a free field. A free field is never experienced in practice, however, and sound levels in rooms fall off at less than 6 dB per doubling of distance, although in most rooms there will be a range of distances that approximate a free field (White and Louie 2005: 204). In practice ... sound not only gets quieter but also gets duller as one moves away from a source. The amount of excess attenuation is dependent on the level of impurities and humidity [in the air] and is therefore variable (Howard and Angus 2006: 29).

General Comments

To determine the location of a sound source along the proximal plane of a mix, a listener can focus on the psychoacoustic information inherent in the following components of the source: its sonic envelope (attack, sustain, decay, release); its relative amplitude (rms & peak); its spectral profile; and the broader decay profile of the envelope at large. Included in the envelope of a sound is, of course, implied acoustic information, whether natural (by capture) or induced via signal processing.

To determine the balance amongst instruments, the listener must consider not only relative amplitude levels but signal processing as well. Visualizing the instruments from the intro along the proximity plane places the hi-hats closest to the listener then the kick drum (when it enters at approximately 0:18) and the electric guitar. One might think to place the kick drum closest to the listener, as it reaches the highest peak amplitude of the three instruments in question, but its short duration, and left-panned reverb processing, move it back-left in the proximity plane; and, moreover, as none of the other instruments share its reverberation profile, it resides in its own unique space in any event. The back-right movement of the hi-hats (via delay processing) in conjunction with the kick drum's back-left movement (via reverb processing) work in tandem to create motion and depth. The effect is only enhanced by the opposing motion of the section's highest (hi-hat) and lowest (kick drum: bottom 60 Hz - 80 Hz, attack around 670 Hz, and beater head around 2.6 kHz) frequencies.

VERSE 1 (0:27-1:01)

Foundation: Kick Drum, Guitars 1/1A, Crash Cymbal & Snare Drum

Uninterrupted audition of the introduction and verse 1 can cause the hi-hat to appear as if it carries the foundation of the mix, when, in fact, the kick drum has moved to beats 2 and 4 to establish a backbeat. Guitars 1 & 1A play straight sixteenth note arpeggios, repeated every two beats; and the bass, hi-hats, and palm-muted guitar (guitar 2), provide the requisite counter rhythm to maintain the (reggae) upbeat rhythmic pattern.

As verse 1 begins, the kick drum moves to the center of the horizontal plane, and its previous reverb processing is jarringly muted. The limited room information now included in the kick drum track capture indicates a close-mic scenario during tracking. Note that the snare rattle heard section introduction no longer sounds. The kick drum remains unprocessed and panned to the center position throughout verse 1, moving closer to the listener along the proximity plane than before. A comparison of the dry kick drum of verse 1, with the processed kick drum from the song's introduction, demonstrates one of the roles signal-processing can play in defining mix function.

Guitar tracks 1 and 1A play identical parts, but are panned to opposite sides of the stereo spectrum. Their timbres are slightly different, indicating different signal paths (possibly the result of alternate pickup selection, a different guitar altogether, a different amplifier, a different microphone, different equalization settings, etc.), but both parts receive the same modulated delay processing, namely, chorus.[32] Instead of a flanger's "whooshing" sound, chorus exhibits detectible pitch modulations. In this case, the use of modulation is subtle but discernible, and it softens the attack transients of guitar tracks 1 and 1A, thus moving them slightly back along the proximity plane. The effect also gives the guitar tracks a modest back-and-forth motion along the proximity plane, as the processor cycles through the modulations defined by its rate and depth controls. The softening effect of the chorus in conjunction with the panning placement reduces the guitars' potential for masking the lead vocal tracks. Guitar tracks 1 and 1A contain room information that indicates a microphone placement far enough back from the amplifier speakers to include room reflections caused by the performance; this additional reverb helps situate the guitars in their own space and contributes to their ability to "sit" behind the vocal tracks. By design, these mixing and tracking decisions "wedge" a space in the center of the horizontal plane for the lead vocal tracks to occupy.

The crash cymbal appears exclusively at the extreme right of the horizontal plane, and only ever on beats 2 or 4, effectively reinforcing the backbeat. The relative amplitude of the crash is quite low compared to the rest of the drum kit, which, sets it back along the proximity plane and allows it to re-texture and accentuate the backbeat. Appearing only with the crash cymbal, the snare drum sees limited use compared to its traditional role in the rock genre, namely, as the anchor point for beats 2 and 4. The

---

[32] Having two identical guitar parts panned to opposite sides of the horizontal plane is conventional practice in pop/rock record production. Each track normally receives slightly different equalization processing but in this case both guitar tracks sound as though they have been played by the same instrument with different pickup settings: the left channel with the bridge pickup and the right with the neck pickup. The guitar heard in this recording is most likely a Fender Telecaster.

snare's timbre is momentarily unmasked at the end of the verse (0:58-0:59), when drummer Stewart Copeland strikes the snare slightly out of time with the kick and crash cymbal.

Counter-Rhythm: Guitar 2, Hi-Hats, & Bass

Guitar 2 exhibits the effects of masking so much so that the track can go completely unnoticed. Similar in timbre to guitars 1 and 1A, guitar 2 fights for space not only in the frequency domain but rhythmically as well. With guitars 1 and 1A occupying every 16th note, little room exists for guitar 2 to eke out a place in the mix. Panned to approximately 11 o'clock, guitar 2 recedes easily into the mix due to its limited sonic envelope and indiscernible spatial information that, were it audible, may have situated the track in a slightly clearer acoustic space. The palm-muted performance contributes a percussive quality that indistinctly textures and accents the counter-rhythm.

The hi-hats to do not appear to retain the slight delay processing found in the introduction. It is entirely possible that the guitar in the right channel masks the space created by this delay line, though the delay bus could just as easily have been muted during mixing. Due to the majority of the hi-hat energy residing in the upper bands of the frequency spectrum, and from the space created around the hats by room information in the recording, they remain distinct in the busier mix of verse 1. Some horizontal movement is created as the hats pan slightly left of center at 0:43-0:46 and again at 0:52-0:54.

Working with the hi-hats and guitar 2 as counter-rhythm to the foundation, Sting's bass track strongly accentuates each up-beat, helping to both establish and maintain the reggae feel of the track at large. The bass track's lowest frequency content sits at approximately 80 Hz, with further resonance around 135 Hz, and with its percussive attack components sound at roughly 1.2 kHz. Panned to the center of the horizontal plane, the bass occupies the lower register of the frequency spectrum along with the kick drum, although they do not play at the same time (a non-standard practice in pop/rock). The fact that the bass and kick play rhythmically opposite each other avoids potential masking issues in the kick/bass relationship. Often recorded through a D.I. box (direct injection box), as seems to be the case in this instance, the bass contains no discernible room information from the recording process, which renders it completely dry (forward) and stable in the mix.[33]

Lead: Vocal 1 & Vocal 2

A combination of two vocal tracks take on the lead role during the verse. Vocal 1, which is higher in register, has a completely different sonic quality in comparison with vocal 2 (lower register). However, the blend between the two vocal tracks (both sung by Sting) creates the illusion of a unified lead. The effects of masking have been put to creative use to blend the timbre of the two vocal tracks in this case, and bussed reverb

---

[33] **Direct Injection Box**: The direct [injection] box is an adapter to allow connecting an instrument pickup, instrument preamplifier, or power amplifier directly to the mic or line input of a recording or sound reinforcement console. This avoids having to use a microphone for acoustic pickup, or offers a different sound quality. Electric guitars, basses, and keyboards are typical candidates for this treatment ... (White and Louie 2005: 111).

processing (with vocal 1 exciting the processor more readily) puts both tracks in the same apparent physical space which adds to the impression of a single lead.

The sonic characteristics of vocal 2 indicate a close microphone capture at the tracking stage. The track contains little room information as the low amplitude of the vocal delivery, and its close proximity to the microphone capsule, minimizes room reflections. Vocal 1, in contrast, includes obvious ambient information due to the greater amplitude of the performance and the distance between the performer and microphone. Vocal 1's room sound is particularly noticeable when the highest/loudest notes sound at 0:40-0:43, 0:50-0:52, and 0:57-1:01. Vocal 2's lack of ambience, or dryness, should technically move it forward in the mix, however, due to the mixing engineer's control over general amplitude levels by track, and the subdued performance, vocal 2's reduced amplitude and sonic characteristics allow it to act as an underlying texture to the dominant vocal track (vocal 1). Bussed reverb processing supplements the cohesion between vocal tracks as it places both vocal performances (and their captured ambience) within a shared space. The controlled width of the reverb processing reveals the behaviour of a stereo reverb bus panned to approximately 10 o'clock on the left channel and 2 o'clock on the right. The bussed reverb not only creates a common space for the lead tracks to occupy but limits the effects of masking by preventing the vocal tracks from competing with the guitars at the edges of the horizontal plane.

General Comments

Substantial changes occur in the mix at the start of verse 1, which help to maintain listener interest and assist in "sectionalizing" verse 1 from the introduction. Of particular note is the rhythmic shift which occurs as a result of mix edits. The rapid and smooth fade-in of guitars 1 and 1A reveals a probable editing "cover up," as beat one completely disappears, creating an odd transition from the intro into the verse. The ability to edit recorded performances is one of the hallmarks of recording practice and, as such, many recordings will yield audible imperfections to the attentive listener. An obvious punch-in at 0:41 on vocal 2 during the elongated vowel sound of the word "sky" illustrates this point.[34]

CHORUS 1 (1:02-1:19)

Foundation: Kick Drum & Bass

Playing together for the first time in the mix, the kick and bass fulfill their traditional role as foundational mix elements. The bass and kick accent each downbeat, and compete for space at the low-end of the frequency spectrum, on beats 2 and 4. An apparent boost applied to the bass range might be the result of doubling, equalization, or the blend of a D.I. and a miked speaker cabinet, but is more likely the effect of modulation processing combined with phase interference. Inserting a phase or chorus

[34] Punch In: The precise control of the onset of tape recording on one or more tracks in the midst of an already existing recording. For instance, a fluffed word or phrase by an announcer can be corrected by listening to the playback and punching in at the exact moment. Punching in requires considerable skill and also a tape recorder designed for it. Today, it is naturally much easier to do this on a digital editor (White and Louie 2005: 314).

effect on the bass track effectively "thickens" the bass part, that is, it widens its position in the mix. Adding chorus to a bass track to achieve such an effect is relatively standard practice in rock/pop record production, having reached widespread prominence in the late 1970s. The slight "warble" indicative of chorusing is audible on the bass track, in my opinion. The increase of amplitude and low-frequency content in the bass track causes it to mask the majority of the kick drum's lower frequencies, leaving the percussive sound of the beater head and some low-mid frequency information to deliver the remaining timbre of the kick. The increase in "size" of the bass guitar elevates its prominence in the mix and sets it just behind the vocals in the proximity plane.

Counter-Rhythm: Hi-Hats, Guitar 1, Guitar 2

The hi-hat track remains relatively unchanged other than an increase in amplitude which brings out slightly more room resonance, and creates a somewhat brighter spectral image. The entire mix boosts in amplitude, so the increase in level of the hi-hats is congruent with the overall amplitude increase of the various elements in the section. The amplitude increase of the section brings everything forward along the proximity plane.

Guitar 1, panned right, plays on the upbeats in typical reggae fashion and the generous use of reverb further alludes to the genre. Along with the spring reverb (generated most likely by the guitar amplifier) the sound of the tracking room imparts its sonic characteristics on the track. The location of the guitar in the horizontal plane, along with the spring reverb and room reflections, place the guitar in its own space and keep it well out of the way of the other mix elements. The less pronounced attack envelope of guitar 2, together with a lower amplitude than many of the other mix elements, set it further back along the proximity plane. Guitar 2 is processed with a moderate chorusing effect which softens its attack and somewhat destabilizes its proximal location.

Lead: Vocal 1 & Vocal 2

The overall amplitude increase of the chorus section brightens the vocal tracks by giving them more apparent presence and energy in the range of 4 kHz and up. Vocal track 2 (lower register delivery) contains more room reflections than it did in the previous section as a result of a different type of microphone technique. Backing off the microphone allows the room reflections to become part of the captured performance, and, combined with the amplitude increase of the section the room ambience becomes more detectable. Both vocal tracks have the same, or very similar, reverb processing as in the previous section.

VERSE 2 (1:20-1:54)

Section to Section Comparative Analysis

At this point in the analysis, more useful insight can be garnered from a comparison of song sections (e.g. verse 1 with verse 2) than by repeating the same analytic-descriptive format as above. For example, an overall change in the use of spatial processing differentiates verse 1 from verse 2. The *lack* of spatial processing on the guitar tracks instantly distinguishes the sonic character of verse 2. A direct a/b comparison of verse 1 and verse 2 demonstrates the cumulative effect equalization, panning, and reverb processing have on proximity. The unprocessed guitars of verse 2 occupy more of a central location than their verse 1 counterparts and without reverb processing move the feel of the entire mix forward. The vocal tracks maintain the same kind of processing they receive in verse 1 but, without the reverb wash and delay tails on the hi-hats and guitars, the effect on the mix is entirely different. The vocals now sit in a sonic space that is in absolute contrast with the rest of the mix.

Chorus 2 & 3 (1:55-2:12 & 2:57-3:14)

The mix consistency from chorus to chorus is virtually identical barring negligible differences in performance. The uniformity from chorus to chorus makes sense, I suppose, from a commercial perspective or, perhaps, the unchanging chorus acts as the foundation for the mix at a macro level, that is, it provides a point of departure for the other sections to interact and contrast with.

BRIDGE/SOLO (2:13-2:56)

Mix Overview

The signal processing on the kick drum is reminiscent of verse 1. The expansive reverb moves the kick back in the mix and creates space for the lead guitar to fill, the latter also being processed to include a unique reverberation profile. The amplitude of the lead guitar, combined with reverb processing, moves the track back along the proximity plane; the track is, as such, spatially situated in a distinct manner, which establishes its "lead" role. The rhythm guitar part (guitar 1) from verse 1 returns during the bridge, although this time only a single track (panned right) appears. The track, heavily processed with reverb and with a relatively low amplitude compared to the other mix elements, sends it well back along the proximity plane. Beginning at 2:29 a combination of kick, snare, and crash cymbal reinforce beats 2 and 4 in the same way as at the end of verse 1.

The bridge section introduces one new sonic element, a synth pad, which enters at 2:21 and sustains as one long note until the end of the bridge at 2:56. The synthesizer pad adds a layer of texture by masking/blending the different instruments that share its component frequencies. Panned wide across the stereo spectrum it is most audible at the extreme left and right sides of the horizontal plane.

OUTTRO (3:15-4:12)

Mix Overview

Similar to the bridge, the outtro serves to recall several elements, sounds, and tracks from the entirety of the recording (a mix "recapitulation," as it were). The outtro sees the return of the muted electric guitar that is heavily masked in verse 1, this time out in the open until masked by the lead guitar track. Also reappearing is verse 1 and 2's rhythm guitar part, this time panned only to the right and with reverb processing approximately half as wet as verse 1's reverb. The lead guitar's sonic character is the same as found in the bridge. It occupies its own defined space which allows it to stand out in the mix. Interacting with the lead guitar in the same horizontal location is a vocal ad lib track. Its relatively low amplitude and off-mic delivery causes it to fight for space with the lead guitar.

As the amplitude of the outtro attenuates towards infinity, which begins at approximately 3:49, reverb processing occurs across the entire mix and progressively increases in wetness. The overall reduction in amplitude combined with the increase of the reverb processor's output multiplies the distancing effect of the fade-out. As the energy of the mix dissipates, it moves farther away from the listener along the proximity plane. The reverb processing causes the dwindling signal to appear as if it is leaving a physical space.

SUMMARY

The preceding analysis focuses explicitly on delay, modulation and reverb processing, and their combined ramifications vis-a-vis the horizontal and proximity planes of a single mix. I hope it is clear by now that I do not assert this analytic focus provides a comprehensive explanation of the broader musical meanings of "Bring On The Night" (1979), nor that it is necessary that every analysis of that track include the kinds of information I offer. That said, my analysis does provide what is the first comprehensive accounting for the musical functions of delay, modulation and reverb processing on this track, which I also argue is crucial for understanding those elements of the track's reception which recordists can, indeed, fix, namely, spatial positioning. Moreover, I hope it is clear that such positioning remains a crucial musical concern for recordists, and that these concerns play a significant role in shaping the final sound of records. I now turn my attention to considering future directions in which this research may be taken to generalize its scholarly appeal.

CHAPTER FIVE

CONCLUSIONS AND FUTURE RESEARCH

In this thesis I have presented an analytic model for examining the musical functions of delay and modulation processing. In so doing, I believe that I have effectively challenged many prevailing assumptions in academe about what, specifically, constitutes "musical practice," by focusing analytic attention on musical procedures and terms reserved for recordists that, until very recently, have only registered in research as extra-musical technologizations of "live" exchange, if at all (Hodgson 2010). I provided an academic context for my model in Chapter One, while, in Chapter Two, I elucidated a number of basic psychoacoustic assumptions on which my model follows, and I explained the technical basis of delay and modulation processing. In Chapter Three, I examined common ways that recordists use the psychoacoustic principles I examine in Chapter Two; and, in so doing, I provided practical information concerning delay and modulation processing which analysts require to follow the application of my model in Chapter Four. I will now briefly consider future research possibilities for generalizing, and expanding upon, what I have thus far presented.

To begin with, I believe that my study fruitfully complicates an underlying assumption about reception which guides its present situation in scholarly research on Recording Practice, that is, that it is inherently polysemantic. While I fail to see any controversy in asserting the fundamental polysemism of listener interpretations of recorded musical communications, many aspects of reception are indeed fixed through the conscious application of numerous psychoacoustic principles during the record making process. Most relevant for my purposes right now are psychoacoustic cues concerning space and spatial location. Delay and modulation processing, and reverb processing (a species of delay), all process signal so it bears psychoacoustic cues concerning location via the listening position a mix constructs. Just as it is a physical impossibility to hear anything above roughly 22 kHz, so, too, is it impossible to hear a dry signal as "closer" than another, with the same amplitude and pitch, processed to bear an RT lasting more than, say, 3 seconds.

My model might, then, be useful for nuancing current models of reception to include physiological responses to sound. Recordists are aware of physiology, and use delay and modulation processing, among other tools, to orchestrate the physiological responses listeners experience during playback. In fact, I would argue it is crucial to do this if analysts are to heed the musical particulars of recorded musical communications, and to provide a more *comprehensive* understanding of musical meaning making vis-a-vis Recording Practice. After all, listener attention is *directed* before they begin to fashion a personally situated interpretation of recorded sound.

This said, future work on the limitations of my model is also necessary. Numerous technological mediations inhere in the reproduction of recorded sound, many of which easily interfere in the transmission of recorded musical communications. My model cannot conform to the so-called "hypodermic model" of musical meaning "in operational and practical fact," as Marshall McLuhan (1968: 15) once wrote. Speakers fail, program equalizations slant psychoacoustic information into bizarre and incoherent formulations, stereo phase and cross-talk ensures distortion of a provided stereo image, and so on. There is also the issue of variations in listener hearing capacity. Nonetheless, these listener interventions are most often anticipated by recordists and "managed" vis-a-vis mixing and mastering.

The analysis I present in Chapter Four is focused explicitly on delay, modulation and reverb processing, and their combined ramifications vis-a-vis the horizontal and proximity planes of a single mix. As I noted at the conclusion of that chapter, I certainly do not assert that this analysis provides a comprehensive explanation of the broader musical meanings of "Bring On The Night" (1979), nor that analysts of that track must examine the kinds of musical information I note to provide a useful accounting for the song's meanings. That said, my analysis does provide what is the first comprehensive accounting for the musical functions of delay, modulation and reverb processing on this track, which I also argue is crucial for understanding those elements of the track's reception which recordists can, indeed, fix, namely, spatial positioning. More tracks would need to be examined to generalize my model to encompass Recording Practice at large. Future research will need to be done, then, to ascertain whether delay and modulation processing is a prerequisite component of recorded musical communications, or singular techniques which recordists invoke as artistic whim and need dictate.

Clearly my research can be taken in numerous directions, and some refinement will be required before this is done. That said, I believe I have indeed contributed to current academic research on Recording Practice, reception of records, and musical meaning making in general, through the model I propose. It is my hope that future researchers will build on the model I present, examining further its generalizability, and its general utility, as an analytic focus for scholars of popular music practice.

WORKS CONSULTED AND CITED

Alluri, V., and P. Toiviainen. 2010. Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre. *Music Perception* 27: 223-241.

Ballou, Glen. 2005. *Handbook for Sound Engineers*. New York: Focal Press.

Beauchamp, James W. 2007. *Analysis, Synthesis, and Perception of Musical Sounds: The Sound of Music.* Urbana: Springer.

Blake, Andrew. 2009. Recording Practices and the Role of the Producer. In *Cambridge Companion to Recorded Music,* eds. Nicholas Cook, Eric Clarke, Daniel Leech-Wilkinson, and John Rink, 36-53. Cambridge: Cambridge University Press.

Blesser, Barry. 2001. An Interdisciplinary Synthesis of Reverberation Viewpoints. *Journal of the Audio Engineering Soc*iety 49.10: 867-903.

Blum, Stephen. 1992. Analysis of Musical Style. ed. Helen Myers. *Ethnomusicology: An Introduction.* New York: Norton.

Braasch, Jonas. 2005. A Binaural Model to Predict Positions and Extensions of Spacial Images Created with Standard Sound Recording Techniques. *Audio Engineering Society, Conference Paper 6610. Presented at the 119th Convention.* New York, October 7-10.

Brackett, David. 1995. *Interpreting Popular Music.* Cambridge: Cambridge University Press.

Brunner, Stefan, Hans-Joachim, and Stefam Weinzierl. 2007. On the Audibility of Comb Filter Distortions. *Proceedings of the 122nd AES Convention, Berlin, May 2007, Paper Number 7047.* Available at: http://www.aes.org/ (accessed May 1, 2010).

Case, Alexander. 2007. *Sound FX: Unlocking the Creative Potential of Recording Studio Effects.* Oxford: Focal Press.

Cook, Nicholas, Eric Clarke, Daniel Leech-Wilkinson, and John Rink, eds. 2009. *The Cambridge Companion to Recorded Music.* Cambridge: Cambridge University Press.

Corey, Jason. 2010. *Audio Production and Critical Listening.* Oxford: Focal Press.

Covach, John. 1997. We Won't Get Fooled Again: Rock Music and Musical Analysis. *In Theory Only* 13: 119:141.

Doyle, Peter. 2004. From 'My Blue Heaven' to 'Race with the Devil': Echo, Reverb and (Dis)ordered Space in Early Popular Music Recording. *Popular Music* 23: 31-49.

__________. 2005. *Echo and Reverb: Fabricating Space in Popular Music Recording 1900-1960*. Middletown, CT: Wesleyan University Press.

Eargle, John, and Ron Streicher. 1990. Acoustical Perspective in Commercial Two Channel Stereophonic Recording. *Proceedings of the 8th International AES Conference, JBL, May 1990, Paper Number 8-020.* Available at: http://aes.org/ (accessed May 1, 2010).

Everest, Alton F. 2001. *The Master Handbook of Acoustics*, 4th Ed. New York: McGraw-Hill.

Fales, Cornelia. 2002. The Paradox of Timbre. *Ethnomusicology* 46: 56-95.

__________. 2006. Sound: A Simple Introduction to its Production, Perception, and Representation. Available at: *http://www.indiana.edu/~savail/ workingpapers/ sound.html* (accessed May 1, 2010).

Fast, Susan. 2000. *In the Houses of the Holy: Led Zeppelin and the Power of Rock Music.* New York: Oxford University Press.

Fastl, Hugo, and Eberhard Zwicker. 2007. *Psychoacoustics: Facts and Models.* Berlin: Springer.

Fisher, John Andrew. 1998. Rock 'n' Recording: The Ontological Complexity of Rock Music. In *Musical Worlds: New Directions in the Philosophy of Music,* ed. Philip Alperson, 109-123. Pennsylvania: University of Pennsylvania Press.

Gallagher, Mitch. 2007. *Acoustic Design for the Home Studio*. New Jersey: Thomson Course Technology.

Gibson, Bill. 2002. *Equalizers, Reverbs and Delays.* Michigan: Pro Audio Press.

Gibson, David. 2005. *The Art of Mixing,* Second Edition. Boston: Artist Pro Publishing.

Gracyk, Theodore. 1996. *Rhythm and Noise: An Aesthetics of Rock.* London: Duke University Press.

Haas, Michael. 2009. Broadening Horizons: 'Performance' in the Studio. In *Cambridge Companion to Recorded Music.* eds. Nicholas Cook, Eric Clarke, Daniel Leech-Wilkinson, and John Rink, 59-62. Cambridge: Cambridge University Press.

Horning, Susan Schmidt. 2002. *Chasing Sound: The Culture and Technology of RecordingStudios in America, 1877-1977.* Ph.D., Case Western Reserve University.

Hodgson, Jay. 2006. *Navigating the Network of Recording Practice—Towards an Ecology of the Record Medium.* Ph.D., University of Alberta.

__________. 2010. *Understanding Records: A Field Guide to Recording Practice.* New York: Continuum.

Howard, David, and Jamie Angus. 2006. *Acoustics and Psychoacoustics*, 3rd Ed. New York: Focal Press.

Huber, David Miles. 2009. *Modern Recording Techniques,* 7th Ed. New York: Focal Press.

Izhaki, Roey. 2008. *Mixing Audio: Concepts, Practices And Tools.* Oxford. Focal Press.

Jones Steve. 1990. The Intro and the Outtro: Technology and Popular Music Practice. *Popular Music and Society* 14: 1-6.

Kaczmarek, Marcin, Cecylia Szmal, and Ryszrd Tomczyk. 1999. Influence of Sound Effects on Sound Quality. *Proceedings of the 106th AES Convention, Wroclaw, May 1999, Paper Number 4902.* Available at: http://www.aes.org/ (accessed May 1, 2010).

Kahrs, Mark and Karlheinz Brandenburg eds. 2003. *Applications of Digital Signal Processing to Audio and Acoustics.* Norwell: Kluwer Academic Publishers.

Katz, Mark. 2004. *Capturing Sound: How Technology Has Changed Music*. Berkeley: University of California Press.

Kealey, Edward R. 1990. From Craft to Art: the Case of Sound Mixers and Popular Music. In *On Record: Rock, Pop, and the Written Word*, ed. Simon Frith and Andrew Goodwin, 207-220. London: Routledge.

Kent, Nick. 1994. *The Dark Stuff: Selected Writings on Rock Music*. London: Da Capo Press.

Kuzmich, Natalie. 2007. Aesthetics: What Does It Mean?. *The Canadian Music Educator* 48: 4-6.

Lacasse, Serge. 2000. Voice and Sound Processing: Examples of Mise en Scene of Voice in Recorded Rock Music. *Popular Musicology Online* 5: Available at: http://www.popular-musicology-online.com/issues/05/lacasse.html (accessed May 1, 2010).

Levitin, Daniel. 2006. *This Is Your Brain on Music: The Science of a Human Obsession.* New York: Plume.

Lewis, Amanda. 2010. *Towards a Model for Analyzing Microphone Practice on Rock Records*. M.A. thesis. The University of Western Ontario. London, Ont.: Faculty of Graduate Studies, The University of Western Ontario. Print.

Lopez-Poveda, Enrique A., Alan R. Palmer, and Ray Meddis, eds. 2010. *The Neurophysiological Bases of Auditory Perception.* Nottingham: Springer.

McLuhan, Marshall. 1968. *The Medium is the Massage: An Inventory of Effects*. New York: Gingko Press.

Middleton, Richard. 1993. Popular Music Analysis and Musicology: Bridging the Gap. *Popular Music* 12: 177-190.

Montgomery, D. 2008. Echo and Reverb: Fabricating Space in Popular Music Recording 1900-1960. *Popular Music and Society* 31: 697-699.

Moore, F. Allan. 2001. *Rock, the Primary Text: Developing a Musicology of Rock,* 2nd Ed. Burlington: Ashgate.

___________. September 2005. Beyond a Musicology of Production. *Paper presented at the conference Art of Record Production, London College of Music*. Available at: http://www.artofrecordproduction.com/content/view/26/52/ (accessed May 1, 2010).

Moorefield, Virgil. 2005. *The Producer As Composer: Shaping The Sounds of Popular Music*. Cambridge: MIT University Press.

Morley, Neil. 2009. Phasing. Harmony Central. Available at:http://www.harmony-central.com/Effects/Articles/Phase_Shifting (accessed May 1, 2010).

Moylan, William. 1987. A Systematic Method for the Aural Analysis of Sound Sources in Audio Reproduction/Reinforcement, Communications, and Musical Contexts. *Proceedings of the 83rd AES Convention, Lowell, October 1987, Paper Number 2514.* Available at: http://www.aes.org/ (accessed May 1, 2010).

____________. 2007. *Understanding and Crafting the Mix: The Art of Recording, Second Edition.* Amsterdam and London: Focal Press.

Normandeau, Robert. 2009. Timbre Spatialisation: The Medium Is the Space. *Organised Sound* 14: 277-285.

Owsinski, Bobby. 1999. *The Mixing Engineer's Handbook.* Auburn Hills: Intertec.

Porcello, Thomas. 2007. Three Contributions to the 'Sonic Turn.' *Current Musicology* 83: 153-166.

Rumsey, Francis, and Tim McCormick. 2009. *Sound and Recording, 6th Ed.* Oxford: Focal Press.

Ryan, Kevin and Brian Kehew. 2006. *Recording the Beatles: The Studio Equipment and Techniques Used to Create Their Classic Albums.* New York: Curvebender Publishing.

Schmidt-Horning, Susan. 2004. Engineering the Performance: Recording Engineers, Tacit Knowledge and the Art of Controlling Sound. *Social Studies of Science.* 34.5: 703-731.

Sethares, William A. 2007. *Rhythm and Transforms.* Madison: Springer.

Stavrou, Michael P. 2003. *Mixing With Your Mind.* Mosman: Flux Research Pty. Ltd.

Sterne, Jonathan. 2003. *The Audible Past: The Cultural Origins of Sound Reproduction.* Durham: Duke University Press.

Talbot-Smith, Michael, ed. 1999. *Audio Engineer's Reference Book.* Oxford: Focal Press.

____________. 2002. *Sound Engineering Explained.* New York: Focal Press.

Tankel, Jonathan D. 1990. The Practice of Recording Music: Remixing as Recoding. *Journal of Communication* 40: 34-46.

Théberge, Paul. 1995. What's That Sound? Listening To Popular Music, Revisited. *Popular Music: Style and Identity*, ed. W. Straw, S. Johnson, R. Sullivan and P. Friedlander, 275-284. Montreal: Centre for Research on Canadian Cultural Industries and Institutions.

____________. 1997. *Any Sound You Can Imagine: Making Music/Consuming Technology*. Hanover, NH: Wesleyan Press.

Toft, Robert. 2010a. Hits and Misses: Crafting a Pop Single for the Top-40 Market in the 1960s. *Popular Music* 29: 267-81.

____________. 2010b. *Hits and Misses: Crafting Top 40 Singles, 1963-1971*. New York: Continuum.

Toole, Floyd E. 2008. *Sound Reproduction: Loudspeakers and Rooms.* Oxford. Focal Press.

Warner, Timothy. 2003. *Pop Music Technology and Creativity: Trevor Horn and the Digital Revolution*. Aldershot: Ashgate.

White, Glenn D. and Gary J. Louie. 2005. *The Audio Dictionary*, 3rd Ed. Seattle: University of Washington Press.

White, Paul. February 2009. Creating A Sense Of Depth In Your Mix: Mix Processing Techniques. *Sound On Sound*. Available at: http://www.soundonsound.com (accessed May 1, 2010).

Wurtzler, Steve. 1992. *Sound Theory Sound Practice*. ed. Rick Altman. New York: Routledge.

Zagorski-Thomas, Simon. 2008. The Musicology of Record Production. *20th Century Music* 4 (2). Available at: http://www.zagorski-thomas.com/ (accessed May 1, 2010).

_______. 2009. The Medium In The Message: Phonographic Staging Techniques that Utilize the Sonic Characteristics of Reproduction Media. *Journal of the Art of Record Production.* Available at: http://www.artofrecordproduction.com/content/view/26/52/(accessed May 1, 2010).

Zak, Albin J. 2001. *The Poetics of Rock: Cutting Tracks, Making Records.* Berkeley. University of California Press.

DISCOGRAPHY

The Police, Regatta De Blanc (A&M: 1979) - Electronic Release
Radiohead, Hail To The Thief (Parlophone: 2002) - Electronic Release