Audio Processing

Michail Exarchos (a.k.a. Stereo Mike) and Simon Zagorski-Thomas

Introduction

This chapter initiates the discussion of post-production through an exploration of the 'state of the art' in both the practice and our theoretical understanding of audio processing, echoing the underlying pursuit of the handbook stated in its introduction – that is, the bridging of dichotomies between the theoretical and the practical (the 'how' and the 'why') in/behind record production processes and outcomes. In some ways, this is also reflected in the tone and examples offered by the two authors, contributing both practice-based/phonographic illustrations, as well as surveys of relevant theorisations from other disciplines. In attempting to provide a theoretical map/ping between musical-sonic objectives, technological/processual actualisation, and pursued aesthetics (personal/stylistic sonic signatures in record production), we examine four domains of sonic characteristics, and ensuing tools and processes devised for their manipulation – namely "pitch", "amplitude", "spectrum and timbre", and "time and space" (processing).

Our discussion of *pitch* processing and associated tools explores the first historical attempts at de-associating the domains of pitch and time via, initially, innovations in the design of analogue devices, which quickly problematises the notion of neat sonic categories when looked upon from the lens of practice; we then examine how the science behind digital pitch-processing devices and algorithms takes advantage of the functional interrelationship between pitch and delay at the heart of their effective operation. The *amplitude* section explores this notion further, suggesting that perhaps (almost) all mixing processes/workflows can be seen as a function of amplitude processing; the implication, however, being that spectral/timbral and dynamic/temporal perceptual effects cannot be understood but as artefacts of processes (ab)using aspects of amplitude (over different domains, such as relative balance, panorama, spectra and time), which – in turn – highlights shortcomings in the discipline's consistency of the current state of our definitions and categorisations. *Spectrum And Timbre* starts from this very issue to bring together understandings from related areas (such as music and audio psychology, plugin design, ecological perception, semantics, production pedagogy, ethnomusicology and popular music studies), demonstrating how embodied knowledge and critical listening can be further informed enriching practice; although the section commences by offering residualist and frequency-specific definitions of timbre, it eventually illustrates through a merging of semantic and plugin-design understandings how our current (software) mixing tools have come closer to a "bundling" of simplified timbral controls that reflect perception rather than causality. Finally, the *Time And Space* section provides a time-based definition of spatial effects as used in record production, illustrating the surreal complexity of staging artefacts in contemporary (popular) music-making, to then underline the gap that exists between technical and musicological theories of production aesthetics.

Pitch

At the heart of contemporary pitch manipulation lies the process of pitch shifting and its close relationship with delay. Although, delay-related processing will be dealt with in more detail in the *Time And Space* section below, the focus here will remain on processes dealing with, and perceived as, forms of pitch manipulation in record production. A number of tools and applications borne out of the creative exploitation of pitch shifting have been created serving record production needs, from pitch correction, through to harmonising, enhancing, and real-time, expressive textural applications. Interestingly, as pitch processing technologies have developed, a cyclic thread can be observed from early efforts attempting to avoid inadvertent "glitch" or granular artefacts as a result of pitch-shifting processes, through to Autotune's self-proclaimed "transparent" pitch-correction, and back to much contemporary pitch manipulation, celebrating and bringing to the fore the textural artefacts infused through – extreme – pitch processing. In other words, we can map the history of pitch processing applications and technologies against inadvertent (caused by technological limitations) and/or conscious (stylistically driven, aesthetically pursued) textural traces. Of course, exploring this causality links pitch processing also to spectral phenomena, which are the direct subject of the *Spectrum And Timbre* section below; but it does make thematic sense to break the processing analysis down to pragmatic units, mirroring real-world record production applications.

To better understand the relationship between pitch processing and the time domain, it is worth exploring some of the earliest attempts at pitch and time manipulation. Much of the literature dealing with the history of pitch manipulation starts off with the Eventide H910 Harmonizer, deemed as 'the first commercially available pitch correction device' (Arcticman4200, 2018), yet one that came with the trade-off of substantial digital artefacts (Owsinski, no date). But as Costello (2010) points out, a 1960s AES article (Marlens, 1966) locates some of the earliest patents relating to pitch and time manipulation back to analogue devices from the 1920s. Marshall informs us that the driver for such innovation had been communication and not record production, offering the 1967 Eltro Information Rate Changer – specifically marketed for "speech compression" – as the earliest commercially available analogue pitch-time changer evolving from these earlier patents:

> The ability to change a recorded sound's pitch independently of its playback rate had its origins not in the realm of music technology, but in efforts to time-compress signals for faster communication … In short, for the earliest "pitch-time correction" technologies, the pitch itself was largely a secondary concern, of interest primarily because it was desirable for the sake of intelligibility to pitch-change time-altered sounds into a more normal-sounding frequency range. (Marshall, 2014)

This relationship between pitch processing and the time domain is not just a functional one – it sits at the core of pitch processing theory and application. Focusing on pitch shifting as key for related digital processes, Izhaki provides a useful definition that exemplifies the relationship:

> A pitch shifter alters the pitch of the input signal without altering its duration ... If we zoom into the recording of a human voice, we identify repeating patterns that only vary over time …. A pitch shifter identifies these patterns, then shrinks or stretches them into the same time span … The actual process is more complex, but at its core is this rudiment. (Izhaki, 2018, p. 469)

To uncover some of the complexities behind the theory of pitch shifting, we can turn to Case (2012, pp. 249-262) who explains pitch 'recalculation' through 'a bit of math', and demonstrates that, although a source sound would simply be shifted in time were it ran through a fixed digital delay, running it through a *variable* delay actually results in *pitch change* (accelerating delay lowers the pitch, whilst decreasing delay raises it). Crucially, he highlights two important theoretical problems: firstly, the manipulation of source content spanning the duration of whole songs would necessitate devices able to produce delay times unthinkable before access to substantial computer-based processing power; secondly, an infinitely varying delay would result in phenomena akin to manipulating analogue playback devices (ibid.). It is worth citing Case at some length here to illustrate the practical implications and how these are dealt with in digital devices:

> Pitch-shifting signal processors differentiate themselves from tape speed tricks in their clever solving of this problem. Digital delays can be manipulated to always increase, but also to reset themselves … It (is) the *rate* of change of the delay that (leads) to pitch shifting, not the *absolute* delay time itself ...
>
> It is a problem solved by clever software engineers who find ways to make this inaudible. Older pitch shifters "glitched" as they tried to return to the original delay time. Today, those glitches are mostly overcome by intense signal processing. Software algorithms can evaluate the audio and find a strategic time to reset the delay time, applying cross fades to smooth things out. (Case, 2012, p. 253)

However, the "glitch" effects perceptible in early digital devices such as the H910 were, of course, put to creative use by record producers. Notable Harmonizer highlights include: Jimmy Page's steel-drum effect on Led Zeppelin's 'Moby Dick/Bonzo's Montreux' (*Led Zeppelin*, 1990); Kevin Killen's detuning of snares on U2's *War* (1983) and *The Unforgettable Fire* (1984); AC/DC's *Back in Black* (1980), featuring Tony Platt's infamous snare, vocal and guitar riff "fattening"; Tony Visconti's metallic snare on 'Breaking Glass' from David Bowie's *Low* (1977); Eddie Van Halen's use of two H910s for much of his ultra-wide 1980s guitar tones, for example, on *5150* (1986); Laurie Anderson's studio and live vocal/instrumental processing as of the late 1970s; and Aphex Twin's use of four units of the later H949 model on the Grammy-winning *Syro* (2014) (Bain, 2017; Eventide Audio, 2016a, 2016b).

Case (2012, p. 253) above, notably, differentiates between older pitch shifters that 'glitched' and today's 'intense signal processing' (carried out via computer software), which facilitates a high degree of transparency. Fast-forwarding to 1997, Antares Auto-Tune's initial – and continuing – impact has been very much a result of achieving this textural transparency, which allowed, arguably, *discreet* pitch correction. Provenzano explains:

> Older analog voice manipulation technologies such as vocoders and talk boxes can bind the output of the larynx to an equal-tempered scale, but there is a cost (or a benefit): the timbre of the voice gets mixed up and reworked, sometimes beyond recognition. Auto-Tune, by contrast, when used not as an overt effect but as a pitch-correction tool, is not timbrally expensive … (Provenzano, 2018, p. 163)

Provenzano here focuses in on the creative and sonic implications of a pitch processing tool that can be deployed for pitch-correction only, leaving the vocalist's (or instrumentalist's) timbral and (non-pitch-related) emotional aspects of the performance intact. This is a processing objective mirrored in all of Antares's promotional and instructional literature since their first Auto-Tune release, highlighting that their pitch correction takes place 'without distortion or artefacts, while preserving all of the expressive nuance of the original performance' (Antares, 2017).

In much of her chapter, 'Auto-Tune, labor, and the pop-music voice', Provenzano (2018) delineates between this functional objective, which allows producers and artists to focus in on *untamed* performance not restricted by pursuits of tuning perfection, and the "overt" use of Auto-Tune as an effect that does incur a *timbral cost*. But it is important to question at which point does our auditory perception transcend from pitch-only to textural (timbral) appreciation? Or, in other words, what is the threshold between pitch correction and timbral effect, and what are the key variables responsible for crossing over from one type of manipulation to the other? To begin answering this we need to briefly review the functionality of Auto-Tune, and other pitch-correction/manipulation software that have been produced since its release – namely Celemony Melodyne, Waves Tune and cases of DAW-specific tuning functionality, such as Logic Pro X's internal Flex Pitch, and Ableton Live's Warp modes / Transpose function.

In their latest user manual for Auto-Tune Realtime (for UAD), Antares identify the settings necessary to achieve the 'Auto-Tune vocal effect', referring to the "overt" effect exemplified by records such as Cher's *Believe* (1998) and T-Pain's *Epiphany* (2007):

> The Auto-Tune Effect is what is technically known as "pitch quantization". Instead of allowing all the small variations in pitch and gradual transitions between notes that are a normal part of the human singing voice, the Auto-Tune Effect limits each note to its exact target pitch and forces instantaneous transitions between notes.
>
> There are basically two elements to producing the Auto-Tune Effect:
>
> 1. Set the Retune Speed to zero.
>
> 2. Choose the correct scale. (Antares, 2017)

It is worth clarifying at this point that the word "auto" in Auto-Tune refers to one of its two main operating modes, the automatic one, which is typically opted for when correcting performances that are generally quite close in tuning to an intended scale. By selecting the respective scale and setting the 'Retune Speed', 'Correction Style', 'Natural Vibrato' and 'Humanize' parameters to values rendering naturalistic results – or, conversely, extreme ones to achieve the "Auto-Tune Vocal Effect" (see figure 21.1 below) – a speedy pitch correction process can be carried out, which may work for the majority of the source content. But in cases where the automatic pitch correction renders some unwanted results (pitch quantisation defaulting to unintended notes of the chosen scale due to largely inaccurate performance), Auto-Tune's Graphic Mode is more apt (and challenging to operate), allowing increased editing/manual accuracy over individual notes and utterances.

Figure 21.1. The Antares Auto-Tune Realtime UAD plugin window showing settings used by one of the authors on the lead rap voice for a recent Trap remix: a softer take on the hard "Auto-Tune Voice Effect" has been achieved (courtesy of the less than maximum Retune Speed setting), and one note of the F# minor scale selected has been bypassed to facilitate the "auto" mode without graphic intervention.

It is in comparison to Auto-Tune's Graphic Mode that Celemony's Melodyne (winner of the Technical Grammy in 2012) has brought creative advantages to the table since its first public viewing at the NAMM show in 2001. Melodyne takes the graphical interface possibilities further and expands the functionality – in later releases – to polyphonic pitch correction and manipulation. In a comparative review for Auto-Tune's version 5 and Melodyne's first plugin version, Walden (2007) points out that Melodyne's 'power takes it beyond the role of a purely corrective processor and into that of a powerful creative tool'. The "powerful" artefact-free algorithm has, as a result, enabled creative manipulation beyond simple pitch correction or hard pitch quantisation and many practitioners have been empowered by its intuitive interface, using it to create everything from Harmonizer-style pseudo-ADT (Anderton, 2018), through to artificial backing harmonies, and – notably also – Skrillex *Bangarang*-style (2011) "glitch" vocals (Computer Music Magazine, 2014).

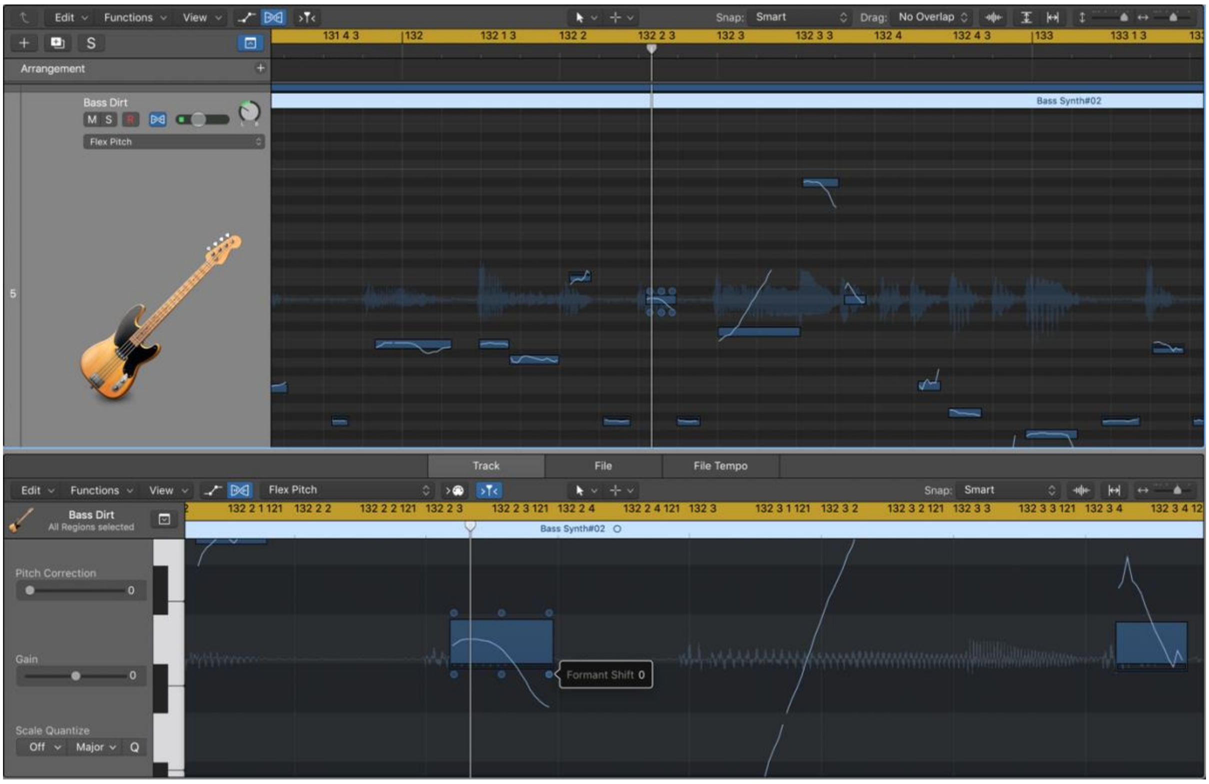

Figure 21.2. Flex Pitch mode enabled on a distorted bass guitar track in Logic Pro X (10.4.1), zooming in on both its Workspace and – the more detailed – Editor views. The six nodes visible on the latter provide access to parameters such as Pitch Drift, Vibrato and Formant Shift.

By 2013, Apple had brought many of Melodyne's features to Logic Pro X in the form of its Flex Pitch functionality (Breen, 2013; Kahn, 2013), providing direct access to advanced processing of multiple pitch parameters inside a DAW (see figure 21.2 above), while a few years prior, Waves had provided a 'mutant child of *Auto-Tune* and Celemony's *Melodyne*' with Waves Tune (Godfrey 2006). Ableton Live's powerful elastic engine, courtesy of their flexible Warp modes algorithms (see figure 21.3 below), arguably became its unique selling point, initially for easy and sonically transparent pitch- and time-stretching; but, eventually, for the glitchy and granular sound design explored by artists such as Flume (for example, on *Skin*, 2016), when its 'granular resynthesis' engine is pushed to extremes (McFarlane, 2017). Although there are arguments for or against both the ease and sound of the now many available algorithm alternatives in pitch processing practice (see, for example, Vandeviver, 2018), it would be fair to say that the current state of the art is certainly at a stage where the user can dial in anything from transparent/functional control or correction, through to convincing pseudo-performative layers, or onto expressive/experimental real-time artefacts – benefitting from the elasticity now available as a direct result of increased processing power.

This brings us back full circle to our timbral thread and its relationship to pitch manipulation. According to Moore (2012): 'Timbre is concerned with the harmonic relationships, phase relationships and the overall volume contour or envelope of the sound', who, further, cites Moylan (2007) suggesting 'that dynamic envelope; spectral content and spectral envelope are all crucial in the perception of timbre'. Although – as the *Spectrum And Timbre* section will show below – the study of timbre is problematised by questions of definition, it can certainly be argued then that pitch manipulation, even at its most discreet, results in textural artefacts of different shades on the scale of perceptibility. At this point we would suggest that it is necessary to further research the timbral (envelope, amplitude, spectral, etc.) ramifications of pitch processing more systematically.

Figure 21.3. Ableton Live's Clip View illustrating a number of available Warp modes and the Transpose function, which are often combined and pushed to extreme settings live or via automation in pursuit of granular effects.

Amplitude

In a similar vein, it is almost impossible to single out amplitude processing in a record/post production context from spectral and time-based causality and ramifications. It is interesting that the literature on amplitude effects and psychoacoustics weaves analytical perspectives that juxtapose notions of amplitude, perceived loudness, the time domain and spectral or harmonic distortion to meaningfully discuss our perception of amplitude in musical sound. Case (2012), for example, breaks down his analysis of amplitude effects into chapters dealing first with distortion, equalisation, dynamics (compression, limiting, expansion and gating), and then volume. His distortion chapter 'leaves the time axis well alone and focuses on the sometimes accidental and sometimes deliberate manipulation of the amplitude axis' (ibid., p. 89). Howard and Angus (2009, pp. 91-2) explain that:

> Although the perceived loudness of an acoustic sound is related to its amplitude, there is not a simple one-to-one functional relationship. As a psychoacoustic effect it is affected by both the context and nature of the sound…
>
> The pressure amplitude of a sound wave does not directly relate to its perceived loudness. In fact it is possible for a sound wave with a larger pressure amplitude to sound quieter than a sound wave with a lower pressure amplitude. 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.

It follows then that we need to consider the various approaches to amplitude processing in record production from the perspectives of creative application (processing tools, practice), pursued effect (aesthetics, perception), *and* their interrelationship. A helpful concept to consider here may be Stavrou's (2003, p. 51) self-proclaimed guiding mixing principle of generating 'Maximum Illusion with Minimum Voltage', highlighting the psychoacoustic oxymoron that exists between abused causality (in record production practice) and sonic effect (in perception). The remainder of this section will look at cases of amplitude processing practiced in a contemporary record production context that pursue a range of loudness perception effects.

Of course, from a post-production perspective, one has to start with faders – and before them, historically, rotary knobs – which have been used 'for coarse level adjustments' and constitute 'the most straightforward tools in the mixing arsenal' (Izhaki, 2018, p. 183). Although there are differences in fader designs (potentiometers providing level attenuation through resistance, VCA faders allowing voltage-controlled amplification, digital faders determining sample multiplication), functionally, they all enable the balancing of levels for individual elements in a mix even though they – arguably by design – inspire turning things up (ibid., pp. 183-187). Due to this, one of the mix engineer's primary concerns becomes "gain-staging": the management of relative levels between different stages of a signal path, allowing sufficient headroom for further level adjustment while maximising signal-to-noise ratios (Davis, 2013; Houghton, 2013). Sufficient headroom in optimum gain-staging parlance technically relates to the avoidance of distortion and noise artefacts but, as Case's (2012) thematic breakdown of amplitude effects demonstrates, distortion of amplitude should very much be appreciated as a processing artefact stemming from the creative application of headroom *abuse*. A classic, if extreme, example of amplitude processing from discography can be heard in the harmonically distorted kicks and snares of countless drum n' bass, hip-hop and EDM records (check, for example, the exposed beat intro to 'Ghosts n' Stuff' from *For Lack of a Better Name* by deadmau5, 2009), achieved through hard "clipping": the process of 'abusing the analogue stages of an A-D converter' resulting in the generation of 'odd harmonics' (Houghton, 2015); an effect initially carried out through the abuse of converters on digital samplers or audio interfaces, later via digital clipping within DAWs operating at lower resolutions than today (it is actually quite hard to reach clipping point within contemporary DAWs running at high-bit floating architectures – see Thornton, 2010), and – currently – typically via third-party clipping plugins. What is interesting about harmonic distortion achieved through hard clipping is that, not only does the process result in increased loudness perception; it is also a case of amplitude processing bringing about a form of *synthesis* – the generation of new harmonics, typically only over the transient phase of a percussive source, with both dynamic and textural implications. This can also be viewed as a fundamentally stylistically-driven sonic signature, with post-production implications for loudness perception, yet again blurring the theoretical delineation of amplitude and timbral processing.

Equalisers are very much part of the same debate, as they are effectively tools that allow amplitude manipulation over a specific part of the frequency spectrum (there are many types of equalisers allowing and achieving different types of spectral manipulation and effect, respectively, covered by a variety of technical handbooks such as some of the ones cited here – see, for example, Case, 2012; Davis, 2013; Izhaki, 2018), but resulting in altered *timbre*. Following the same logic, independent amplitude manipulation of the signal levels of two or more channels (typically via panoramic potentiometers) – mapped to different speakers – results in imaging or spatial aural illusions of placement or movement. Finally, amplitude processing over time pertains to our use of dynamic range processors, such as compressors, limiters, gates, expanders and duckers. Perceiving these tools and applications under the conceptual theme of time-based amplitude processing is helpful, as is their delineation between our processing foci on the micro and macro temporal domain. It would not be a stretch to suggest that most mixing processes revolve around four notions of amplitude processing:

- as relative level balancing;
- as panning;
- over specific areas of the frequency spectrum (equalisation); and
- over time (dynamics);

leaving only ambient/spatial processing as a separate case (although there are many amplitude-related functions to this aspect as well, but perhaps it cannot be as neatly explained solely in relation to amplitude as either parameter or type of processing). This notion may explain the unavoidable crossover between our theoretical delineations of different processing categories, and why they are so entangled in practice. We would therefore suggest that there is a need to complicate our theoretical mapping of the interrelationship between physical/acoustic/technological causality and psychoacoustic perception effects, taking into account the semantically incomplete attempts at categorisation and definition of sonic phenomena in record production musicology.

Spectrum And Timbre

The study of timbre has been plagued by questions of definition. These range from the notion that all sound is only frequency spectrum over time i.e. there is not anything but timbre, to a kind of residualist approach in music theory – that timbre is what is left over in the process of orchestrating and arranging after pitch, rhythm, dynamics, harmony and form have been accounted for i.e. it differentiates instrument types and delineates playing techniques (excluding dynamics) such as pizzicato, marcato, con legno, sul ponticello / tasto, con sordini etc. In light of this and of similar problems around the notion of spatial and time-based effects, these last two sections will include brief surveys of the ways in which these subjects are theorised in other disciplines as well as exploring how they are approached through the techniques of record production. Hopefully this will provide further insights from both sides of the poeisis/esthesis fence.

In the world of audio and music psychology, a good deal of work has developed around the problems of identification and categorisation (e.g. Bregman, 1994; McAdams *et al.*, 1995; Peeters *et al.*, 2011). How do we distinguish between a trumpet and violin while at the same time being able to categorise the whole range of different timbres that a trumpeter can produce as being ‘the same’? Some of this work is based on the notion of ‘feature extraction’ – of looking for statistical correlations between sounds that are identified as in some way ‘the same’. Feature extraction is, of course, at the heart of the notion of emulation in digital effects and plugin design and research (e.g. Yeh, Abel and Smith, 2007; Zölzer, 2011; Paiva *et al.*, 2012) and that has become hugely important in the commercial world of record production as the notion of timbre has become more and more associated with vintage products and the sonic signatures of star mixers rather than with principles or theory. Porcello (2004, p. 747) has discussed the way that sound engineers speak about sound and that one of the key ways that this happens is through association, in particular by indexically invoking industry professionals or production technologies. This has now become embodied in the design and marketing techniques of product designers and the economics of the sector is driving research towards emulation rather than innovation. Of course, that is not stopping practitioners from using these emulative tools in innovative ways.

In parallel to this detailed empirical and practical work, there are also several strands of work that are approaching the notion of timbre from the perspective of what it means to people. In the world electroacoustic music, flowing from the idea of creating and listening to sound in an abstract way, various composer-researchers (e.g. Schaeffer, 1977; Smalley, 1986, 1997; Wishart, 1997) have explored both how our perception and interpretation of music is tied to the cause of a sound (Clarke, 2005) and how we can seek to break those ties. These discussions relate both to the nature of the 'thing' that makes the sound (by vibrating) and the nature of the energy that causes it to vibrate. Without much direct cross-over in either direction, this has also been an important factor in work within ethnomusicology and popular music studies (Fales, 2002, 2005; Berger and Fales, 2005; Fales, 2017; Fink, Latour and Wallmark, 2018; Zagorski-Thomas, 2018). Of course, these approaches are reflected in the first chapter of this Handbook that used the categories of agents (human and non-human), energy, space, context and media (the representational systems used to produce and disseminate the sounds) as a framework for analyzing recorded music. This semantic approach to audio (who did what? Where, when and why?) is also reflected in recent developments in plugin design. This is seen, for example, in the Waves CLA series (branded on producer / mixer Chris Lord-Alge) where there are parameters labelled with terms such as spank, roar, bark and honk, or the Infected Mushroom Pusher (branded on the Infected Mushroom EDM duo) which includes punch, push, body and magic.

Of course, this semantic approach takes us away from tools that were perceived to be controlling only one parameter and towards bundled effects that affect dynamics / envelope, frequency / spectrum (including distortion), ambience / reverberation and other effects such as chorus and delay. Elsewhere in this Handbook both Meynell and Zak have discussed the ways in which the Pultec EQ1A affects the dynamics of the signal as well as the frequency content and a lot of the emulation plugins we have been describing are dynamic processors which also add spectral coloration and distortion in some desirable manner.

The key texts used in production pedagogy, on the other hand, (e.g. Owsinski, 1999; Case, 2007; Izhaki, 2008) treat spectral processing (equalization and distortion) in generic terms. The technical details of both equalization and distortion are covered in detail and a series of established approaches are also discussed. And although there are variations in the detail of how those mentioned above and other authors (e.g. Savage, 2011; Mynett, 2013) describe the reasons for equalization, there are three fundamental approaches:

- To make a particular feature of a sound stand out more e.g. to exaggerate the high frequency consonant sounds (t, k, p, s etc.) in a vocal to improve intelligibility
- To remove some unwanted feature like noise or an unpleasant resonance
- To improve clarity and prevent masking by ensuring different sounds in a mix are not competing in the same frequency range

One of the reasons for making a particular frequency range stand out more, and this might be done by equalization of distortion, is to change the perceived energy expenditure of the activity that is causing the noise. This might be quite subtle, in the sense of bringing out energy that is already there – heaviness or lightness in a gesture for example – or it might be about creating quite a surreal cartoon of high energy through distortion – although distortion can signify the vibration of a degraded artefact as well as a higher level of energy in making it vibrate. In any event, there are a range of views about the possibility of specifying good and bad generic equalization settings. We all have our 'go to' approaches for specific contexts – Phil Harding outlines one for popular music earlier in this book – but it is also true that the 'go to' approach is a starting point from which to use critical listening and decide what else needs to be done.

Time And Space

It may seem odd to put time and space under the same heading but the idea is that time-domain effects – in contrast to pitch, amplitude and spectral effects – related to most of the techniques that were concerned with spatial processing: delay and reverberation most notably. They are techniques that are based, in the acoustic world at least, on reflections and therefore on delayed copies of some original 'direct' signal. Given that definition, it also makes sense to place phasing, flanging and chorusing in the same box although they all also involve either different frequencies being delayed by different amounts or by the delays being caused by different and varying playback speeds of the 'copied' signal which causes both time and pitch differences. An addition level of time-domain effects is the world of looped samples, audio quantizing and time warping but these are dealt with by Anne Danielsen elsewhere in the Handbook and so will not be covered here.

In the same way that we dealt with spectrum and timbre from a range of perspectives, we will start by discussing the ways in which the world of audio and music psychology has addressed this issue. Not surprisingly, it has been focused on the way that our perceptual and cognitive systems deal with spatial audio in the 'real world' (e.g. Howard and Angus, 1996; Moore, 1999; Begault, Wenzel and Anderson, 2001; Rumsey, 2001) and, given that the vast majority of the funding for research in this area is focused on virtual reality and the gaming industry, the work on technologies of production, recording and emulation are similarly focused on realism (Abel *et al.*, 2006; Maxwell, 2007; Schlemmer, 2011). And although, as with Abel's work on emulating spring reverbs, some of the 'realism' is concerned with 'real' vintage technologies, one set of workflows that is emerging is to select very realistic reverbs and to combine them in a very surreal *sonic cartoon.*

Once again, these empirical and practical approaches exist in parallel with a range of strands that explore acoustic space in terms of its meaning. Reaching back to Edward Hall's (1966) work on proxemics, the anthropology of architecture, we can find a range of work in sound studies, social history and cultural theory that runs in a similar vein (e.g. Schmidt-Horning, 2012; Born, 2013; Dibben, 2013; Sterne, 2015b, 2015a; Revill, 2016). It is concerned with psychological and cultural notions of space, intimacy and the many ways in which the way we exist in the natural and built environments influence our identities and interactions. Being culturally tied to the concert hall more than to commercially distributed recording formats has allowed and encouraged the world of electroacoustic music to embrace the many 3-D formats of spatial audio while also approaching the notion of space in ways that are simultaneously more surreal and more literal (Zotkin, Duraiswami and Davis, 2004; Smalley, 2007; Kendall, 2010). The use of surround sound and 3-D formats has encouraged composers to surround and immerse the listener with the sound source but the virtual spaces in which the sound sources are staged are often very literal.

Within the world of popular music, on the other hand, the case is very often reversed: the instrumentation is often very conventional while the staging is often highly surreal, with multiple spaces of different sizes superimposed on each other. Although the complexities of staging in popular music have been discussed much more than in other musical traditions (e.g. Moore, 1992; Doyle, 2006; Camilleri, 2010; Dockwray and Moore, 2010; Moore, 2012; Dibben, 2013; Zagorski-Thomas, 2014), these very common practices of multiple parallel and often conflicting spatial cues need to be addressed in more detail. Indeed, it is often the case that small, bright room ambience is used purely as a timbral and dynamic effect, thickening up and lengthening the tone of drums for example, while additional reverbs and delays are considered to be doing the spatial work.

Stereo spatial effects involving time-based processing (delay and reverb) are very seldom realistic – even binaural recordings are seldom dummy-head and the levels of 'artificiality' can be altered in many different ways:

- Configuring two microphones differently than two ears on a head: e.g. pointing two directional mics towards the sound source or spacing them wider apart than a head
- Using more microphones than we have ears and then mixing them down to a stereo master afterwards
- Using artificial or schematic reverb that can then be processed separately from the 'dry' original sound
- Superimposing more than one reverb on top of each other, combining delay and reverb or changing the perceived performance environment in different sections of a piece of music

Much like the question of equalization and distortion, the established literature focuses (Owsinski, 1999; Case, 2007; Moylan, 2007; Izhaki, 2008; Minchella, 2018) on the theory of the generic technologies rather than the sonic specifics of particular units or spaces. There is also quite extensive reference to the practices and techniques of UK and US popular music from the 1960s to the 1990s and the ways in which reverberation was often delayed, equalized or filtered to prevent masking. Indeed, whether it through interviews with key engineers (Owsinski, 1999) or reference to landmark recordings (Case, 2007; Moylan, 2007; Izhaki, 2008), the literature harks back to the apprentice model of learning by example from a master with relatively little in terms of detailed analysis of the justification or musical reasons for particular approaches. Thus Izhaki explains delay and reverb coupling in the following terms:

> Blending a delay with a reverb is known to result in more impressive effect than having only one of them. One way to look at it is as if the delay enhances the reverb (Izhaki, 2008, p. 452)

And Case on unreal space:

> …reverb devices are also used for synthesizing a reverberant character that may not exist in nature. Wishing to enhance the music with lushness or some other form of beauty and held only to a standard that it "sound good", an engineer might dial in settings on a reverb device that violate the physics of room acoustics…Such a reverb can be made to sound glorious coming out of loudspeakers. (Case, 2007, p. 309)

This reflects one of the key historical problems with pedagogy in this area in that the scientific theory behind the devices is something to be learned but that the aesthetics of how they should be applied in musical circumstances are a matter of long-term experience and learning 'hints and tips' from the masters about what they did. There is very little connect between the technical theory and a musicological theory of production aesthetics. To be fair, this is a disconnect that exists across the whole world of music education but it is a disconnect that we should be working hard to address.

Conclusion

The chapter's narrative has, on the one hand, followed a linear path informed by its domain-categorisation, but, on the other hand, it has simultaneously challenged it by progressing in its *dynamics* from questions brought about by practice/process, which highlight issues in our theoretical understanding, to *syntheses* of cross-disciplinary theoretical frameworks, which should leave – it is our hope – a circular (non-linear) *resonance* in the reader's mind. The words emphasised here are, of course, purposefully chosen to reflect the introduction's music production metaphors, at the same time illustrating that a completely textual interpretation of sonic phenomena may be limited (i.e. the musical analogy is more than a playful metaphor – it functions as a structural interpretative tool as well). The chapter's sections therefore connect, contrast and interact in their illustration of how specific domain classifications are problematised by both audio processing tool designs and creative application, identifying gaps in the discipline's use of meaningful definitions/understandings that need to be addressed in the pursuit of a more complete musicological theory of record production aesthetics.

References

Abel, J. S. *et al.* (2006) 'Spring reverb emulation using dispersive allpass filters in a waveguide structure', in. *Audio Engineering Society Convention 121*, Audio Engineering Society.

Anderton, C. (2018) *Double-tracking, harmonizing, and layering: how to record and mix multiple vocals.* Available at: https://reverb.com/news/double-tracking-harmonizing-and-layering-how-to-record-and-mix-multiple-vocals (Accessed: 26 June 2019).

Antares (2017) *Auto-Tune Realtime: User manual.* Antares Audio Technologies.

Arcticman4200 (2018) *The evolution of pitch correction.* Available at: https://davesmusicaladventures.wordpress.com/2018/03/01/the-evolution-of-pitch-correction/ (Accessed: 22 June 2019).

Bain, K. (2017) *How the Eventide Harmonizer expanded the possibilities of sound.* Available at: https://daily.redbullmusicacademy.com/2017/11/eventide-instrumental-instruments (Accessed: 29 June 2019).

Begault, D. R., Wenzel, E. M. and Anderson, M. R. (2001) 'Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source', *Journal of the Audio Engineering Society*, 49(10), pp. 904–916.

Berger, H. M. and Fales, C. (2005) '"Heaviness" in the Perception of Heavy Metal Guitar Timbres: The Match of Perceptual and Acoustic Features Over Time.', in Greene, P. D. and Porcello, T. (eds) *Wired For Sound*. Middletown, CT: Wesleyan University Press, pp. 181–197.

Born, G. (2013) *Music, sound and space: Transformations of public and private experience*. Cambridge University Press.

Bourbon, A. and Zagorski-Thomas, S. (2017) 'The Ecological Approach To Mixing Audio: agency, activity and environment in the process of audio staging', *Journal on the Art of Record Production*, 11. Available at: http://arpjournal.com/the-ecological-approach-to-mixing-audio-agency-activity-and-environment-in-the-process-of-audio-staging/ (Accessed: 10 February 2018).

Breen, C. (2013) *Review: Logic Pro X loses none of its power, gains great new features.* Available at: https://www.macworld.com/article/2044283/logic-pro-x-loses-none-of-its-power-gains-great-new-features.html (Accessed: 8 July 2019).

Bregman, A. S. (1994) *Auditory scene analysis: The perceptual organization of sound*.

Camilleri, L. (2010) 'Shaping Sounds, Shaping Spaces', *Popular Music*, 29(2), pp. 199–211.

Case, A. U. (2007) *Sound FX: Unlocking the Creative Potential of Recording Studio Effects*. Burlington MA: Focal Press.

Clarke, E. F. (2005) *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press, USA.

Computer Music Magazine (2014) *Skrillex-style twisted, glitchy vocals using Melodyne - part 4/10.* Available at: https://www.youtube.com/watch?v=orbroFt1_8E (Accessed: 26 June 2019).

Costello, S. (2010) *Pitch shifters, pre-digital.* Available at: https://valhalladsp.wordpress.com/2010/05/04/pitch-shifters-pre-digital/ (Accessed: 24 June 2019)

Davis, D. (2013) *Sound system engineering.* 4th ed. New York: Focal Press.

Dibben, N. (2013) 'The intimate singing voice: auditory spatial perception and emotion in pop recordings', in Zakharine, D. and Meise, N. (eds) *Electrified Voices: Medial, Socio-Historical and Cultural Aspects of Voice Transfer*. Göttingen: V&R University Press, pp. 107–22.

Dockwray, R. and Moore, A. F. (2010) 'Configuring the sound-box 1965–1972', *Popular Music*, 29(2), pp. 181–197. doi: 10.1017/S0261143010000024.

Doyle, P. (2006) *Echo And Reverb: Fabricating Space In Popular Music Recording, 1900 - 1960*. Wesleyan University Press.

Eltro Brochure (1967) http://www.wendycarlos.com/other/Eltro-1967/Eltro-1967.pdf (2008) (Accessed: 26 June 2019).

Eventide Audio (2016a) *History of the Eventide H910 Harmonizer®*. Available at: https://www.youtube.com/watch?v=977Sri5EcCE&app=desktop (Accessed: 26 June 2019).

Eventide Audio (2016b) *Laurie Anderson and the Eventide H910 Harmonizer (Full Version)*. Available at: https://www.youtube.com/watch?v=6veZObvDCXU&app=desktop (Accessed: 26 June 2019).

Fales, C. (2002) 'The paradox of timbre', *Ethnomusicology*, 46(1), pp. 56–95.

Fales, C. (2005) 'Listening to timbre during the french enlightenment', in. *Proceedings of the Conference on Interdisciplinary Musicology (CIM05), Montréal*.

Fales, C. (2017) *Issues of Timbre: The Inanga Chuchotee*. na.

Fink, R., Latour, M. and Wallmark, Z. (2018) *The Relentless Pursuit of Tone: Timbre in Popular Music*. Oxford University Press.

Godfrey, J. (2006) *Waves Vocal Bundle*. Available at: https://www.soundonsound.com/reviews/waves-vocal-bundle (Accessed: 26 June 2019).

Hall, E. T. (1966) *The Hidden Dimension*. London: Doubleday.

Houghton, M. (2013) *Gain staging in your DAW software: Level Headed.* Available at: https://www.soundonsound.com/techniques/gain-staging-your-daw-software (Accessed: 3 July 2019).

Houghton, M. (2015) *Q. What's the best way to clip my drums?* Available at: https://www.soundonsound.com/sound-advice/q-whats-best-way-clip-my-drums (Accessed: 3 July 2019).

Howard, D. M. and Angus, J. (1996) *Acoustics And Psychoacoustics*. Oxford: Focal Press.

Izhaki, R. (2008) *Mixing Audio: Concepts, Practices and Tools*. Focal Press.

Kahn, Jordan (2013) *Logic Pro X review: Powerful new features & a simplified UI with no compromises for pros.* Available at: https://9to5mac.com/2013/07/26/logic-pro-x-review-powerful-new-features-a-simplified-ui-with-no-compromises-for-pros/ (Accessed: 8 July 2019).

Kendall, G. S. (2010) 'Spatial perception and cognition in multichannel audio for electroacoustic music', *Organised Sound*, 15(3), pp. 228–238.

Marlens, W.S. (1966) 'Duration and frequency alteration', *Journal of the Audio Engineering Society*, 14(2), pp. 132-139.

Marshall, O. (2014) *A brief history of Auto-Tune.* Available at: https://soundstudiesblog.com/2014/04/21/its-about-time-auto-tune/ (Accessed: 23 June 2019).

Maxwell, C. B. (2007) 'Real-time reverb simulation for arbitrary object shapes', in. *Proc. Int. Conf. on Digital Audio Effects, Bordeaux, France*.

McAdams, S. *et al.* (1995) 'Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes', *Psychological research*, 58(3), pp. 177–192.

McFarlane, B. (2017) *Learn the perks of Ableton Live's Warp function.* Available at: https://www.emusician.com/how-to/learn-the-perks-of-ableton-lives-warp-function (Accessed: 27 June 2019).

Minchella, D. (2018) 'The Poietics of Space: The Role and Co-performance of the Spatial Environment in Popular Music', in Bennett, S. and Bates, E. (eds) *Critical Approaches to the Production of Music and Sound*. New York: Bloomsbury Publishing, pp. 41–61.

Moore, A. (2012) 'All buttons in: An investigation into the use of the 1176 FET compressor in popular music production', *Journal on the Art of Record Production*, 6.

Moore, A. F. (1992) *Rock: The Primary Text : Developing a Musicology of Rock (Ashgate Popular and Folk Music Series)*. 2nd edn. Ashgate Pub Ltd.

Moore, A. F. (2012) *Song Means: Analysing and Interpreting Recorded Popular Song*. Farnham: Ashgate Pub Ltd.

Moore, B. C. J. (1999) 'Controversies and Mysteries in Spatial Hearing', in. *Audio Engineering Society 16th International Conference: Spatial Sound Reproduction*, Rovaniemi, Finland: Audio Engineering Society.

Moylan, W. (2007) *Understanding and Crafting the Mix: The Art of Recording*. 2nd edn. Focal Press.

Mynett, M. (2013) *Contemporary Metal Music Production*. PhD. University of Huddersfield.

Owsinski, B. (no date) *The subtle art of pitch correction.* Available at: https://www.uaudio.com/blog/pitch-correction-basics/ (Accessed: 25 June 2019).

Owsinski, B. (Bobby O. (1999) *The Mixing Engineer's Handbook*. 001 edn. Artistpro.

Paiva, R. C. *et al.* (2012) 'Emulation of operational amplifiers and diodes in audio distortion circuits', *IEEE Transactions on Circuits and Systems II: Express Briefs*, 59(10), pp. 688–692.

Peeters, G. *et al.* (2011) 'The timbre toolbox: Extracting audio descriptors from musical signals', *The Journal of the Acoustical Society of America*, 130(5), pp. 2902–2916.

Porcello, T. (2004) 'Speaking of Sound: Language and the Professionalization of Sound-Recording Engineers', *Social Studies of Science*, 34(5), pp. 733–758.

Provenzano, C. (2018) 'Auto-Tune, labor, and the pop-music voice', in Fink, R., Latour, M. and Wallmark, Z. (eds.) *The relentless pursuit of tone: timbre in popular music*. New York: Oxford University Press, pp. 159-181.

Revill, G. (2016) 'How is space made in sound? Spatial mediation, critical phenomenology and the political agency of sound', *Progress in Human Geography*, 40(2), pp. 240–256.

Rumsey, F. (2001) *Spatial Audio*. Focal Press.

Savage, S. (2011) *The Art of Digital Audio Recording: A Practical Guide For Home and Studio*. New York: Oxford University Press, USA.

Schaeffer, P. (1977) *Traité des objets musicaux*. Paris: Le Seuil.

Schlemmer, U. (2011) 'Reverb Design', in. *Pure Data Convention, Weimar-Berlin, Germany*.

Schmidt-Horning, S. (2012) 'The Sounds of Space: Studio as Instrument in the Era of High Fidelity', in Frith, S. and Zagorski-Thomas, S. (eds) *The Art of Record Production: an introductory reader to a new academic field*. Farnham: Ashgate.

Smalley, D. (1986) 'Spectromorphology and Structuring Processes', in Emmerson, S. (ed.) *The Language of Electroacoustic Music*. London: Macmillan, pp. 61–93.

Smalley, D. (1997) 'Spectromorphology: explaining sound-shapes', *Organised Sound*, 2(2), pp. 107–126.

Smalley, D. (2007) 'Space-form and the acousmatic image', *Organised Sound*, 12(1), pp. 35–58.

Stavrou, M.P. (2003) *Mixing with your mind: Closely guarded secrets of sound balance engineering.* Flux Research Pty.

Sterne, J. (2015a) 'Space within space: Artificial reverb and the detachable echo', *Grey Room*, pp. 110–131.

Sterne, J. (2015b) 'The Stereophonic Spaces Of Soundscape', in Théberge, P., Devine, K., and Everrett, T. (eds) *Living Stereo: Histories And Cultures Of Multichannel Sound*. New York: Bloomsbury, pp. 65–83.

Thornton, M. (2010) *Headroom & Master Fader.* Available at: https://www.soundonsound.com/techniques/headroom-master-fader (Accessed: 3 July 2019).

Tompkins, D. (2011) *How to wreck a nice beach: The vocoder from World War II to hip-hop: The machine speaks*. Melville House.

Vandeviver, C. (2018) *Why I prefer Flex Pitch over Melodyne for pitch editing in Logic – expert opinion.* Available at: https://www.pro-tools-expert.com/logic-pro-expert/2018/8/30/why-i-prefer-flex-pitch-over-melodyne-for-pitch-editing-in-logic-expert-opinion (Accessed: 26 June 2019).

Walden, J. (2007) *Auto-Tune vs Melodyne.* Available at: https://www.soundonsound.com/reviews/auto-tune-vs-melodyne (Accessed: 25 June 2019)Wishart, T. (1997) *On Sonic Art*. 2nd Revised edition. Routledge.

Yeh, D. T., Abel, J. S. and Smith, J. O. (2007) 'Simplified, physically-informed models of distortion and overdrive guitar effects pedals', in. *Proc. of the Int. Conf. on Digital Audio Effects (DAFx-07)*, Citeseer, pp. 10–14.

Zagorski-Thomas, S. (2014) 'An Analysis of Space, Gesture and Interaction in Kings of Leon's "Sex On Fire" (2008)', in von Appen, R. et al. (eds) *Twenty-First-Century Pop Music Analyses: Methods, Models, Debates*. Farnham: Ashgate Publishing Limited.

Zagorski-Thomas, S. (2018) 'Timbre As Text: The Cognitive Roots Of Intertextuality', in Lacasse, S. and Burns, L. (eds) *The Pop palimpsest: Intertextuality In Recorded Popular Music*. Ann Arbor, MI: University of Michigan Press, pp. 273–90.

Zölzer, U. (2011) *DAFX: digital audio effects*. John Wiley & Sons.

Zotkin, D. N., Duraiswami, R. and Davis, L. S. (2004) 'Rendering localized spatial audio in a virtual auditory space', *IEEE Transactions on multimedia*, 6(4), pp. 553–564.

Discography

AC/DC (1980) *Back in Black* [vinyl]. Australia: Albert Productions.

Aphex Twin (2014) *Syro* [CD]. UK & Europe: Warp Records.

Cher (1998) *Believe* [CD]. US: Warner Bros. Records.

David Bowie (1977) *Low* [vinyl]. UK: RCA Victor.

Deadmau5 (2009) *For Lack of a Better Name* [CD]. USA & Canada: Ultra Records.

Flume (2016) *Skin* [CD]. US: Mom + Pop.

Laurie Anderson (2001) *Life On A String* [HDCD]. US: Nonesuch.

Led Zeppelin (1990) *Led Zeppelin* [4xCD, RM + Box, Comp]. Europe: Atlantic.

Skrillex (2012) *Bangarang* [CD] US: Big Beat.

T-Pain (2007) *Epiphany* [CD]. US: Jive.

U2 (1983) *War* [vinyl]. UK: Island Records.

U2 (1984) *The Unforgettable Fire* [vinyl]. UK: Island Records.

Van Halen (1986) *5150* [vinyl] US: Warner Bros. Records.