The effect of dynamic range compression on the psychoacoustic quality and loudness of commercial music

inter.noise 2010
13-16 June 2010 . Lisbon . Portugal
noise and sustainability
www.internoise2010.org

The effect of dynamic range compression on the psychoacoustic quality and loudness of commercial music

William Campbell[1], Rob Toulson[2], Justin Paterson[3]

[1] Anglia Ruskin University, Cambridge, UK
{William.Campbell@anglia.ac.uk}
[2] Anglia Ruskin University, Cambridge, UK
{Rob.Toulson@anglia.ac.uk}
[3] London College of Music, Thames Valley University, UK
{Justin.Paterson@tvu.ac.uk }

Abstract

It is common practice for music productions to be mastered with the aim of increasing the perceived loudness for the listener, allowing one record to stand out from another by delivering an immediate impact and intensity. Since the advent of the Compact Disc in 1980, music has increased in RMS level by up to 20dB. This results in many commercial releases being compressed to a dynamic range of 2–3 dB. Initial findings of this study have determined that amplitude compression adversely affects the audio signal with the introduction of audible artifacts such as sudden gain changes, modulation of the noise floor and signal distortion, all of which appear to be related to the onset of listener fatigue.

In this paper, the history and changes in trends with respect to dynamic range are discussed and findings will be presented and evaluated. Initial experimentation, and both the roadmap and challenges for further and wider research are also described and discussed. The key aim of this research is to quantify the effects (both positive and negative) of dynamic range manipulation on the audio signal and subsequent listener experience. A future goal of this study is to ultimately define recommended standards for the dynamic range levels of mastered music in a similar manner to those associated with the film industry.

Keywords: Quality, Psychoacoustics, Music, Loudness, Dynamics.

1 Introduction

In modern popular styles of music, there is a trend to dynamically compress (henceforth, compress) commercial releases in order to enhance its loudness characteristic. There could be a point however, when the negative effects start to outweigh the benefits since by definition, compression is a form of distortion, and over-compression of music can result in perceptible signal degradation. Although there are numerous newspaper, magazine, and online articles addressing the subject, the opinions of the numerous producers, engineers, music critics and columnists have yet to be quantified. Furthermore, the effects (both positive and negative) of compression applied to music have not previously been quantified with respect to perceived signal quality and the subsequent listener experience.

This paper builds upon existing research from the field of hearing psychology, intended for dynamics processing in hearing aids, but specifically discusses how this research might be translated to music. The paper analyses some of the ways in which music might be considered over-compressed and discusses some of the effects. With a focus only on preliminary methodology, this paper perhaps raises more questions than offers explanations, but acts as a roadmap for future work, and more detailed and targeted investigations.

2 History of Compression

The human perception of the loudness of audio, indeed music, can be often be associated with long term (program length) loudness as recently specified by the ATSC [2], based upon the IRU-R BS.1770 standard. This might be related to the average root-mean-square (RMS) power of the audio signal [3], [4]. Complex models of how loudness correlates to RMS power and frequency have been developed in the past by prominent scientists [5], [6], [7], [8]. From 1960 to 1980 the average RMS signal power of commercial music (predominantly on vinyl discs) increased by approximately 4 to 6 dB, but with the advent of the digital Compact Disc (CD) in the early 1980's, and the transition from analogue to digital audio production, commercial music increased in RMS power a further 20 dB between 1980 and 2000 [9], [3]. Many record industry executives dictate that modern popular music productions should have as high an RMS power as possible, driven by the need to 'stand out' in different playback environments by being louder than the competition. This results in many commercial music releases being compressed to a dynamic range of 2–3 dB, as highlighted in previous research by the authors [10]. The dynamic range in this context is defined as the difference between the RMS value and the peak signal value. It is possible to raise the RMS power and perceived loudness of audio using compression and limiting.

An example can be drawn from the rock music genre by mapping the career outputs of Dave Grohl, of bands 'The Foo Fighters' and formerly 'Nirvana' (Figure 1). The RMS values of his albums show a steady increase from *Nevermind* (1991) to *There is Nothing Left to Loose* (1999). The change in the dynamics trend changes course from *One By One* (2002) through *Echoes, Silence, Patience and Grace* (2007) when the Foo Fighters took over as principal producers of their own albums. The loudest tracks on the record are of equal RMS as previous albums, but the records have more diversity and hence some tracks have a lower RMS in recent years.

Contrast in mastering conventions is highlighted in the difference between Guns and Roses production *Chinese Democracy* (2008), which has, at points, a 14 dB dynamic range and the Metallica album *Death Magnetic* (2008), containing tracks compressed to a 2 dB dynamic range (Figure 1). The Mastering Engineer, Ted Jensen wrote in a posting at the (albeit informal) Mettallicabb fan forum,

> *"In this case the mixes were already brick walled before they arrived at my place. Suffice it to say I would never be pushed to overdrive things as far as they are here. Believe me I'm not proud to be associated with this one, and we can only hope that some good will come from this in some form of backlash against volume above all else"* [11].

Brian Moore and colleagues [12], [13], have conducted work in relation to the psychology of hearing and compression to advance the functionality of hearing aids, and it is hypothesized here that the findings of these experiments might be applied directly to music. The greater objective of experiments associated with this project (of which this paper is an early study) is to begin to test the hypothesis that ‘over-compression’ of music might have detrimental effects on the perception of audio quality, contribute to hearing loss, cause listener fatigue, and/or that that the amount of compression has an emotional affect on the listener. Indeed, such work will lead towards a better understanding of the commonly used term, over-compression.

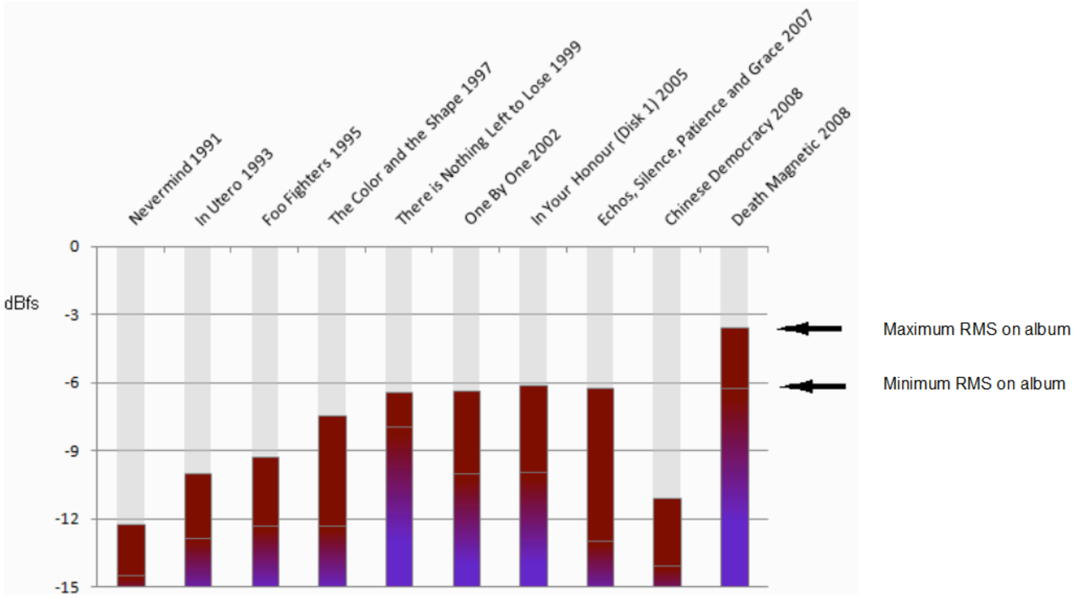

Figure 1 - RMS trend of the early 90’s to present. [10]

3 Compression Experiments for Hearing

3.1 Amplitude compression effects on spoken word audio

Extensive research has been conducted into the perceived loudness and the influence of compression for the advancement of hearing aid design [8], [14]. More recently, Stone et al. have conducted listening tests to analyse the effect of compression when applied to spoken word [13], utilizing two listening tests by means of audio compressed with three different compression ratios (no compression or 1:1, moderate compression or 1.82:1, and severe compression or 10:1; the compression threshold being 3 dB below the RMS level in each case). Both tests asked listeners to identify key phrases or words from passages read by two speakers simultaneously. The tests were conducted as follows: Test one; the two voices were compressed separately then summed, which does not induce cross modulation; Test two; the two voices were summed then compressed (this is how compressed signals are broadcast), which did introduce cross-modulation between the two voices. The tests showed that intelligibility decreased depending on the compression response rate, compression ratio and compressor design as well as with an increased number of compression channels. Sound sources that are normally perceived separately can become perceptually fused after being compressed. Stone et al. state:

“Compression reduces temporal contrast, that is, amplitude variation of the signal envelope over time, the amount of reduction depending on the speed, compression ratio, and design of

the compressor. Multichannel compression also reduces spectral contrast, the amplitude variation across frequency, the amount of reduction depending on number of channels, compression speed, and compression ratio. Loss of both temporal and spectral contrast reduces the information available to the auditory system." [13]

The research performed to date highlights the importance of understanding the effects of compression on audio quality and perceived loudness. Much of the psychoacoustic research to date in this field is limited to pure sinusoids and spoken word, and this kind of analysis has never before been performed using music as the subject audio.

3.2 Applying existing experimental techniques to music

Some prior investigations into the psychoacoustic properties of hearing have been specifically aimed at hearing aid development, which is a key benefit driver for loudness analysis in speech. However, using music as the subject audio brings a whole new range of interest related to culture, creative industries and corporate demands, and the results of this existing research cannot be assumed to hold true for music [10]. The extent to which this does or does not hold true is yet to be quantified.

Similar tests to those defined by Stone et al. [13], have been used to investigate the effects of multichannel compression on music and individual musical instruments. Through signal analysis, effects such as distortion, signal to noise ratio, incidental noise amplification etc, on discrete and mixed musical signals are being investigated. Utilizing and expanding on the methods of Stone et al. [13], the music created for analysis might be used in the future for listening tests to ascertain how compression affects listener perception of loudness and audio quality.

4 Compression Experiment Applied to Music

4.1 Recording (tracking) and mixing

Two methods of recording sample data were considered for this paper, reflecting two different approaches to recording music in a professional studio environment. The first method is live recording; recording instruments played simultaneously in a shared (albeit controlled) acoustic environment. This method is a naturalistic form of recording that accepts microphone spillage rather than attempting to avoid it completely. Another common recording practice is overdubbing, a method of recording that requires the instruments to be recorded at separate times, building up the required layers of a 'live' band. This approach can completely avoid the issues associated with spillage.

Whilst both methods require investigation, in the context of this paper, data collection was limited to the latter method in order to avoid potential (latent) cross modulations of the embedded spilt signals. Such an approach is simpler and potentially more conclusive for the subsequent signal analysis phase of the project, which will form the subject of a future paper.

Tracking of the song *Uncover My Eyes* from the Cambridge (UK) band Bijoumiyo, took place over two days in the Recital Hall at Anglia Ruskin University, produced and engineered by author #1 of this paper and three assistants. Combinations of close and ambient microphones were utilized to provide flexibility during mixing, as is common practice within the recording industry. Great care was taken in the selection and placement of microphones during tracking in order to ensure that the subsequent mixing of the track was already streamlined. As a result, no signal processing, other than stereo panning, was required in either the recording or mixing processes (the latter of which is extremely rare in contemporary industry practice), but this subsequently minimized the creation of artifacts/distortions prior to the application of test compression.

4.2 Mastering

The use of the term 'mastering' for this paper is misleading as only a single part of the typical modern mastering process is performed, i.e. dynamic amplitude compression in the form of limiting, however it is convenient to utilize the expression here to indicate post-mixing manipulation. Here it should be noted that limiting is regarded as an extreme form of compression, where the compression ratio is fixed to a value greater than 10:1 and the threshold level is adjusted to control the amount of compression applied. A 'makeup gain' amplification stage often comes after the compression stage in order to raise signal levels (either RMS or peak) to those comparable with the unprocessed audio.

Two methods of applying compression used in contemporary mastering were utilized in these experiments; discrete signals mixed and then compressed; discrete signals mixed to stems, compressed and mixed (summed without altering signal level) again.

Three levels of compression were applied to the two signal sets, representing heavy, moderate and light compression, all of which being applied using a limiter to better mimic the amount of compression being applied to music referenced in section 2. All adjustments were made to work optimally for signal set A with identical settings applied to signal set B to maintain continuity. Stone et al. applied compression with the threshold set 3 dB below RMS [13], however it was found that applying these settings to music, specifically to stems or discrete signals would alter the mix levels and the output of each experiment would not be comparable. The following are explanations of how the limiter was adjusted:

For light compression, using the Waves L1 software processor, the threshold was set to attenuate the signal as little as possible, simulating limiting being used as a safety device for controlling only the highest peaks in the signal. The release time was set to its fastest setting, exclusive of distortion as perceived by the producer. For moderate compression, the threshold was set to attenuate the signal as little as possible in the very quiet intro of the track (as seen in Figure 4) and to attenuate the signal consistently throughout the rest of the track, but never allowed to stay engaged for more than approximately half a second, thus controlling the overall dynamics of the entire song. Again the release set to its fastest setting without obvious audible distortion.

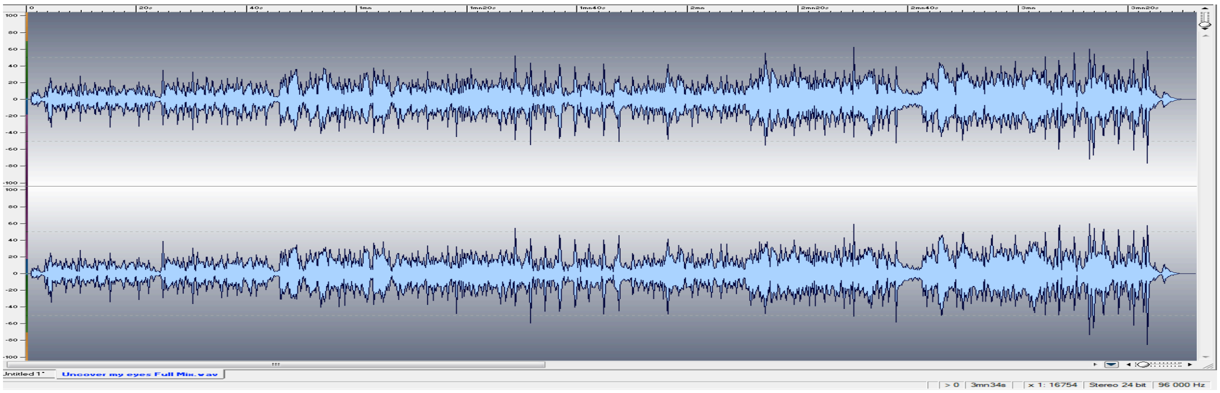

Figure 4 – Waveform of complete song showing low level intro section.

Finally, for heavy compression, the threshold was set to attenuate the signal constantly throughout the entire track, whilst not removing the inherent dynamics completely. Adjusting the threshold to lower levels had no significant effect on the compressor and created an unusable amount of distortion. In fact, in an attempt to reduce the dynamic range to 2 dB, additional supplementary compression was used to no effect other than causing further distortion and reduction of the signals amplitude. The release was set to its fastest setting whilst achieving as little distortion as possible. However, with the heavy amount of limiting involved, distortion was consistently present. Additionally, approximately 3 dB makeup gain was applied to bring the RMS value of the heavily limited track to a comparative level with the light and moderately limited ones.

4.3 Results

The results of analyzing all the tracks after applying the various amounts of limiting show that heavy limiting applied to stems (Set B) results in better dynamic retention than when applied to a full mix (Set A) (Table 1). Heavy limiting on stems is comparable to light compression on the full mix; i.e. both have similar dynamic ranges and both subjectively sound similar. It is probable that the quality of the stems mix can be improved further by adjusting the parameters of the limiters to suit the stems rather than using the settings adjusted for the full mix.

Track/Compression	Dynamic Range	Maximum Peak	RMS
Set A (Full Mix) No compression	17 dB	-1.6	-22.2
Set A (Full Mix) Light compression	11 dB	-8.4	-22.3
Set A (Full Mix) Moderate compression	8 dB	-11.9	-22.6
Set A (Full Mix) Heavy compression	6 dB	-14.8	-22.3
Set B (Stems Mix) No compression	17 dB	-1.6	-22.1
Set B (Stems Mix) Light compression	15 dB	-3.7	-22.3
Set B (Stems Mix) Moderate compression	14 dB	-4.0	-22.3
Set B (Stems Mix) Heavy compression	11 dB	-8.2	-22.3

Table 1 - Limiting measurements

4.4 Discussion of results and opportunities for future work

When comparing like for like i.e. Set B to Set A (e.g. both under heavy compression) from a strictly subjective point of view, Set B, the compressed stems sound better: the drums have more punch and are less distorted, the ostensible noise floor (air conditioning and breathing noises) is lower, there is much more space in the mix, phase distortions are less noticeable in the trumpet and cymbals, and the vocals are less ducked. Each of these artifacts could be subjectively identified in Set A. This would concur with the views of advocates of 'Stem Mastering', a subset of current professional mastering practice, equivalent in its approach.

Due to the stated subjectivity, the greater ongoing study will seek verification through detailed signal analysis & quantification.

Through the use of only limiting with shared temporal responses, possible conclusions here are highly constrained, but this methodology has now been imported in a skeletal form from psychoacoustic studies, and demonstrated itself to be worthy of much more detailed investigation. Such investigation must include detailed cross mapping of:

- Variation of compression ratios – it is anticipated that a matrix of combinations and their quantified effects will be generated for Set B, the stems.
- Variation of temporal settings; a huge number of combinations of attack and release values might be considered. These must again be mapped into a matrix and the various cross modulations compared.
- Both of the above when compression is being applied to peak or RMS values.
- Variation of thresholds, either systematically, or relative to, for example a specified LU reference or RMS.
- All of the above, but subsequently normalized, not just to peak, but also to a specified LU reference or relative to RMS.
- All of the above, but applied as a contemporary producer of multi-track material might, often to individual instruments and channels, as well as sub-groups, yielding an even larger matrix.
- Again, all of the above, but utilizing various spectral separations via multi-band compression.

When all of the dynamic processing is done, all limited signals might be scrutinized using various analytical tools to ascertain and further define the effects of 'over compression'.
All subjective appraisal of such work must be validated through controlled data collection via established clinical survey technique.

5 Conclusions

In conclusion, it has been shown that Moore and Stone's experiments for the advancement of hearing aid development can be applied to music to clearly discernable effect. Whilst compression of sub-groups is a long established practice, and stem mastering has focused on this approach, these early stages of this research project demonstrate the transferability of the methodology, and provide the impetus to collect and collate more elaborate data sets. Stone et al. applied compression with the threshold set 3 dB below RMS [13], and the subsequent evaluation of resultant data was formidable. Applying this to music will yield orders of magnitude more variable, although these might be ameliorated by implementation of only those that sustain musical ‘relevance’. It may be however, that patterns relating to psychoacoustic perception might emerge, enabling a more simplistic distillation for application on a practical level in a number of fields.

Through pursuing this methodology as described above, a new insight into the quantification and perception of long established studio practice will be gained, thus opening the door for its future evolution.

Acknowledgments

We acknowledge the contributions made by the students who assisted during the recording session: Chris Manning, Myles Hill, Lee Rasberry, and thank you to Myles Sanko and the rest of the band, ‘Bijoumiyo’.

References

[1] ITU-R, Draft BS.1770: Algorithms to measure audio programme loudness and true peak level, Geneva, Switzerland, April 2006.

[2] ITU-R, Draft BS.1771: Requirements for loudness and true-peak indicating meters, Geneva, Switzerland, April 2006.

[3] KATZ, B., 2007. Mastering Audio, the Art and Science. Second edn. Canada: Focal Press.

[4] FEY, F., 2009. March for Quality, Studio Magazine, 02/09, 24-29.

[5] FLETCHER, H. and MUNSON, W.A., 1933. Loudness, Its Definition, Measurement and Calculation. *J. Acoust. Soc. Am,* **5**(2), 82-108.

[6] ROBINSON, D.W. and DADSON, R.S. 1956. A re-determination of the equal-loudness relations for pure tones, British J. of Appl. Psychology., 7, 166-181.

[7] HELLMAN, R.P., 1985. Perceived Magnitude of Two-Tone-Noise Complexes: Loudness, Annoyance, and Noisiness. *J.Acoust. Soc. Am,* 77, 1497-1504.

[8] GLASBERG, B.R. and MOORE, B.C.J., 2002. A Model of Loudness Applicable to Time-Varying Sounds. *Journal of the Audio Engineering Society,* 50(5), 331-342.

[9] NEWELL, P., 2000. 7.1 The transition from analogue to digital, Chapter 8 Mastering. *Project Studios, A More Professional Aproach.* First edn. Great Britan: Focal Press, pp. 105-133.

[10] TOULSON, E.R. and CAMPBELL, W., 2009. The life and death of dynamic range: who decides how loud? *The Art of Record Production Conference, Cardiff, Wales,* .

[11] CASHMERE, P., Sept 28 2008, 2008-last update, Metallica Recording Engineer Ted Jensen Craps On Death Magnetic [Homepage of Undercover.com.au], [Online]. Available: http://www.undercover.com.au/News-Story.aspx?id=6423 [Nov, 20, 2009].

[12] MOORE, B.C.J., 2004. An Introduction to the Psychology of Hearing. Fifth edn. Great Britain: Elsevier Academic Press.

[13] STONE, M.A., MOORE, B.C.J., FÜLLGRABE, C. and HINTON, A.C., 2009 July/August. Multichannel Fast-Acting Dynamic Range Compression Hinders Performance By Young Normal-Hearing Listeners In A Two Two-Talker Separation Task. *Journal of the Audio Engineering Society,* 57(7/8), 532-546.

[14] MOORE, B.C.J., VICKERS, D.A., BAER, T. and LAUNER, S., 1999. Factors Affecting the Loudness of Modulated Sounds. *J. Acoust. Soc. Am,* 105, 2757-2772.

Discography

BIJOUMIYO, 2010. Uncover My Eyes. Unpublished.

FOO FIGHTERS, 2007. Echoes, Silence, Patience & Grace. U.S.A., United Kingdom: RCA.

FOO FIGHTERS, 2005. In Your Honor. U.S.A., United Kingdom: RCA, Bertelsmann Music Group.

FOO FIGHTERS, 2002. One by One. U.S.A., United Kingdom: RCA.

FOO FIGHTERS, 1999. There Is Nothing Left to Lose. U.S.A., United Kingdom: RCA.

FOO FIGHTERS, 1997. The Colour and the Shape. U.S.A., United Kingdom: Roswell/Capitol.

FOO FIGHTERS, 1995. Foo Fighters. U.S.A., United Kingdom: Roswell/Capitol.

GUNS N' ROSES, 2008. Chinese Democracy. U.S.A., United Kingdom: Geffen.

METALLICA, 2008. Death Magnetic. U.S.A., United Kingdom: Vertigo, Warner Bros.

NIRVANA, 1993. In Utero. U.S.A., United Kingdom: DGC Records.

NIRVANA, 1991. Nevermind. U.S.A., United Kingdom: DGC Records.