esmuc
ESCOLA SUPERIOR DE MÚSICA DE CATALUNYA

A methodology for the sonic analysis of modern pop music:

Critical listening and its application in four top mixes from 2020

Estudiant: Alejandra Sáenz López

Especialitat: Sonologia

Director: Perfecto Herrera Boyer

Curs: 2021-2022

Vistiplau del director/a del treball

Gràcies al meu tutor, Perfecto Herrera, per la implicació i ajuda al llarg d'aquest treball i durant els quatre anys de grau i a la meva família i amics per acompanyar-me sempre.

Abstract

In recent years, the post-production process of music has gained importance among artists and popular opinion. In a lot of cases, recording, producing and mixing shape the song, bringing out what makes it unique and contributing greatly to the interest and attraction of the music.

My personal interest in the genre and in music analysis, as well as the lack of research on this matter, have brought out the need to seek for a sonic analysis method for pop music, useful for both, interested amateurs and professional musicians. The question I aim to answer is: How can one, in a systematic way, unfold and describe the sonic elements that characterize pop songs? With this intention, I will explore and describe the features that characterize pop music with the help of digital musical analysis tools. I will demonstrate the approach found in four different practical cases and in the end, I hope to be able to reflect on the way that post-production shapes the music.

The work is aimed to serve as a model and a guide to critical listening and sonic analysis based around the book “Recording Analysis: How the Record Shapes the Song", by professor William Moylan, with the intent to simplify the information found in the book, for its level of specification becomes a wall and a tough endeavor for amateur readers. A new analysis method is conformed, from the key aspects of his methodology, and complemented with theoretical information on sound, psychoacoustics and mixing techniques, to achieve general comprehension on sound and its characteristics, to be applied to analyze pop music

The end result is an extensive guide on sonic analysis focused on three main elements we can center on to describe a piece of music: timbre, space and loudness/dynamics. This information is collected via critical listening and contrasted with several digital analysis tools, then structured and compared, to accomplish a general understanding of that piece of music and the contribution that production and mixing make it.

Extracte

En els últims anys, el procés de postproducció de música ha guanyat importància entre els artistes i l'opinió popular. En molts casos, l'enregistrament, producción i mescla exposen el que fa única a la música i contribueixen en gran mesura a l'interès i l'atracció d'aquesta.

El meu interès personal en el gènere pop i en l'anàlisi musical, així com la falta de recerca en aquest tema, han posat de manifest la necessitat de buscar un mètode d'anàlisi sonor per a la música pop, útil tant per a aficionats com per a músics professionals. La pregunta que vull respondre és: Com es poden desenvolupar i descriure, de manera sistemática, els elements sonors que caracteritzen les cançons pop? Amb aquesta intenció, exploraré i descriuré les característiques de la música pop amb l'ajuda d'eines digitals d'anàlisi musical i demostraré el mètode trobat en quatre casos pràctics. Al final, espero poder reflexionar sobre la forma en què la postproducció moldeja la música.

Aquest treball té com a objectiu servir com a model i guia per a l'escolta crítica i l'anàlisi sonor, basat en el llibre "Recording Analysis: How the Record Shapes the Song", del professor William Moylan, amb la intenció de simplificar la informació trobada en el llibre, ja que considero que el seu nivell d'especificació dificulta la comprenssió per als lectors aficionats. Per aquesta raó, he conformat un mètode d'anàlisi, des dels aspectes clau de la seva metodologia, i complementat amb informació teòrica sobre el so, la psicoacústica i les tècniques de mescla, per aconseguir la comprensió general sobre el so i les seves característiques, que després he aplicat per analitzar música pop.
El resultat final és una guia extensa sobre l'anàlisi sonor centrada en tres elements principals sobre els quals podem descriure una peça de música: timbre, espai i sonoritat/dinàmica. Aquesta informació es recopila a través de l'escolta crítica i es contrasta amb diverses eines d'anàlisi digital, després s'estructura i compara, per aconseguir una comprensió general d'aquesta peça de música i entendre com hi contribueixen la producció i la mescla.

Extracto

En los últimos años, el proceso de postproducción de música ha ganado importancia entre los artistas y la opinión popular. En muchos casos, la grabación, producción y mezcla de una canción exponen aquello que la hace única y contribuye en gran medida al interés y la atracción de la misma.

Mi interés personal en el género y en el análisis musical, así como la falta de investigación sobre este tema, han puesto de manifiesto la necesidad de buscar un método de análisis sonoro para música pop, útil tanto para los aficionados como para los músicos profesionales. La pregunta que quiero responder es: ¿Cómo se pueden desarrollar y describir de manera sistemática los elementos sonoros que caracterizan a las canciones pop? Con esta intención, exploraré y describiré las características que particularizan a la música pop con la ayuda de herramientas de análisis musical digital y demostraré el método de análisis encontrado en cuatro casos prácticos. Al final, espero poder reflexionar sobre la forma en que la postproducción moldea la música.

Este trabajo tiene como objetivo servir como modelo y guía para la escucha crítica y el análisis sonoro basado en el libro "Recording Analysis: How the Record Shapes the Song", del profesor William Moylan, con la intención de simplificar la información encontrada en el libro, puesto que considero que su nivel de especificación se convierte en un impedimento para los lectores aficionados. Por esto, he configurado un nuevo método de análisis, desde los aspectos clave de su metodología, y complementado con información teórica sobre el sonido, la psicoacústica y las técnicas de mezcla, para conseguir la comprensión general sobre el sonido y sus características, que luego he aplicado para analizar música pop.

El resultado final es una guía extensa sobre el análisis sonoro centrado en tres elementos principales sobre los cuales podemos describir una pieza de música: timbre, espacio y sonoridad/dinámica. Esta información es recopilada a través de la escucha crítica y contrastada con herramientas de análisis digital, después estructurada y comparada, con el fin de conseguir una comprensión general de esta pieza de música y entender de qué manera contribuyen la producción y la mezcla.

Index

1. Introduction

Pop music is a genre that finds itself in constant stylistic evolution. Changes in society provoke changes in the music and vice versa, making it a flowing genre that moves alongside social trends and is usually representative of the youth and of brief periods of time. Due to this, it demands constant updating of the tools and techniques used for creation, and continuous search of ways with which to surprise the listener and create new trends. In the last decades, technology has made its way into most areas of music making: "Technology has become a 'mode' of music production and consumption: that is, technology has become a precondition for music-making, an important element in the definition of musical sound and style, and a catalyst for musical change" (Blacking, 1977). And at the same time, it has also become, in a lot of cases, a source of inspiration and the cause of creation: "Pop artists and consumers have often used technology in ways unintended by those who manufacture it. In this way, pop practices constantly redefine music technologies through unexpected or alternative uses "(Théberge, 2001).

Despite the fact that technology is key in the making of music, the technical aspects involved in it are usually not given the same appreciation and importance by the media or the popular opinion as the creative aspects (composition, arrangement, performance, etc.). This is probably due to the fact that more specialization is needed to perceive and understand mixing and mastering nuances, and the creative side of the job gets masked under the technical jargon. As a result, this information is usually looked over and the sonic type of analysis is harder to come across in papers or music blogs. This is one of the reasons that conform the motivation for this dissertation, to fulfill this informative gap and put in value the sonic characteristics of today's pop music. I consider it is currently one of the genres that strongly pushes the sonic boundaries. It is constantly exceeding itself, searching for new techniques to achieve the most impact and interest, as well as, adapting to changes in society and the new ways of listening and experiencing music.

The interest I have for this music along with my personal enthusiasm for sound have led me to the desire to dive in and understand the music around me. I am curious to figure out what makes these songs sound interesting to me (and to many of my peers), and unravel what is not heard at a simple listen, eventually incorporating some of the techniques found here into my own future work. I believe this could be interesting for any music fan with a deep interest in the sonic nuances of pop, by giving them the necessary tools to dive into the journey of analysis with the aim to help us better understand our experience of music.

The book "Recording Analysis: How the record shapes the song" (Moylan, 2020) is the main reference for conforming my methodology. It contains highly valuable information and presents a very detailed and thorough structure, although I find it is too extensive and daunting to a newcomer, and therefore it is not easily digestible for amateur listeners. I have taken reference from his lists of analyzable sonic elements and graphical representation ideas (mostly the idea of visualizing several aspects by positioning them in a vertical axis), as well as using his analysis questions to help present conclusions from the analyses. Altogether, I attempt to construct a more accessible analysis method, extracting the elemental aspects of his studies and presenting them in a more visual, dynamic manner. Other references, such as, "Analyzing popular music" (Moore, 2003), or "The Art of Record Production" (Frith & Zagorski-Thomas, 2012), are also valuable sources of information with a different approach to Moylan's. They are composed of short essays by different authors, covering topics on music technology and history. They touch on sonic aspects more briefly and its purpose is not an in-depth analysis of sonic characteristics, but an overview of interesting sonic elements and mixing practices employed in popular songs. "Analysis is inevitably reductive, which is precisely why it's useful" (Moore, 2003). They have been useful to get a hold of different approaches to the production of popular music and the different roles in it. Lastly, the book: "How pros make hits" (Frampton, 2019), contains several short analyses of modern popular songs of the last few years. It is the most "practical case" book and from it I have extracted useful graphic representation ideas, mainly on representation of structure and space, and also it served as a good model for the use of appropriate vocabulary for sound description.

As a result, this paper is divided in four chapters. In chapter one, the motivation, objectives and state of the matter are presented. Chapter two is aimed to give context and historical background on the genre's history, its characteristics and its relationship with technology. In the following chapter, the methodology is proposed, divided in seven sections: biases acknowledgements and suitable analysis conditions, appropriate

description vocabulary, perception aspects to contemplate, digital tools to aid the analysis, graphical representation methods, structure of the analysis and selection of the songs presented. In the final chapter, chapter four, the methodology is applied in four practical cases of 2020 mainstream pop songs.
The dissertation ends with some concluding thoughts, lessons learned, and some ideas to for future work on the development of the analysis method.

2. Music & technology

2.1 Brief history of pop

In order to analyze pop music, first we should state what the term pop music encompasses and where it comes from. Britannica encyclopedia describes it as “any commercially oriented music principally intended to be received and appreciated by a wide audience, generally in literate, technologically advanced societies dominated by urban culture”. On the other hand, The New Grove Dictionary of Music and Musicians identifies it as the music since industrialization, in the 1800, that is most in line with the tastes and interests of the urban middle class. According to these definitions, we could state that the term pop music refers to music that serves as an entertainment for large audiences, mostly young and of urban classes in developed countries. But, how did music become a massive means of entertainment?

Historically, music had been primarily a vehicle for religion and in parallel, popular (non-secular) forms of it emerged. It wasn’t until the creation and popularization of recording and playback devices that music became a massively popular form of entertainment. In the book “The Cambridge Companion to Pop and Rock”, Simon Frith (2001), a British sociomusicologist specialized in popular music culture and professor at University of Edinburgh, gives a clear understanding of how music came to be a mass entertainment by linking it to the ability to store it. From the invention of modern European music notation (XIVth century) to Gutenberg’s press in 1501, and eventually to recording technology (first phonographs and wax cylinders were created by Thomas Edison in 1888), technological developments had a big impact on the consumption and popularization of music as a form of entertainment. “The phonograph meant that public musical performances could now be heard at home; the portable gramophone and transistor radio meant that music moved into the bedroom; the Walkman meant that each individual could make a personal listening tape for use even in public places [...] the industrialization of music, as both a technological and an economic process, describes how music came to be defined as an essentially individual experience ” (Frith, 1987).

During the first half of the 20th century, the music charts consisted of rock and roll (a direct successor of blues, and influenced by the creation of the electric guitar in the early

50s), R&B/soul and jazz songs, with some of the biggest acts being Elvis Presley, Tony Bennett, Les Paul and Mary Ford, The Ronettes, etc. During the 60s, we can see bands such as The Beach Boys and The Beatles started to make their first appearances in the top charts (in 1964 they had their emergence in the Billboard Hot Charts with "I Get Around" and "I Want to Hold Your Hand" respectively). In 1966, the album "Pet Sounds" by The Beach Boys came out and the following year "Sgt. Pepper's Lonely Hearts Club Band" by The Beatles was also released. These albums brought new sonorities, structures, harmonies, etc. which became hugely popular and were unheard before. The focus started to shift from rock to pop, and the phenomenon of pop culture and fandom was induced.

These bands set the precedent and paved the way for the construction of pop as we know it. Some of the musicians that defined the genre during the 70s, 80s and 90s were the Jackson Five, Elton John, Michael Jackson, Prince, Madonna, etc. by incorporating the sound of electronics, synthesizers, sequencers, sampling, etc. in ways never heard before. Their work is still very present in productions today and in the last few years we have seen a tendency to bring back the 70s and 80s sound in modern productions from big artists like Dua Lipa, Bruno Mars, Lady Gaga, etc.

2.2 Pop music and technology

The technological developments during the XXth century had an impact in the composition process too. Multitrack recording[1] was comercially spreading by the end of the 1950s, allowing for a lot more flexibility when mixing and manipulating the recorded material, and new techniques, such as overdubbing[2] (made popular by Les Paul and Mary Ford during the 1950s), were created and are still fundamental today.[3] From Les Paul and

[1] The process of recording separate sound sources, without a defined location in a sound field, in order to combine them to create a cohesive whole

[2] A technique used in audio recording in which audio tracks that have been pre-recorded are then played back and monitored, while simultaneously recording new, doubled tracks onto one or more available tracks. The overdub process can be repeated multiple times.

[3] Other techniques listed in the book "Technological rationalization and the production of popular music" (Theberge, 1989) derived from John Eargle's book "Sound recording" , include: (1) recording in acoustically dry studios; (2) the use of numerous, closely placed microphones to maximize separation; (3) the engineer's participation in the adjustment of musical balances; (4) the introduction of artificial reverberation which could be controlled by the engineer; and (5) careful selection, placement, and balancing of the individual microphones in the stereo array."

Mary Ford, going through the productions by Joe Meek, Phil Spector, Brian Wilson or George Martin, it became clear that “Multitrack recording is not simply a technical process of sound production and reproduction; it is also a compositional process and is thus central to the creation of popular music at the most fundamental level” (Eno, 1983).

Eventually, the evolution of multitrack recording demanded more track recording capacity, and by the 1970s, 24-track recording was available in the biggest studios, allowing engineers to explore: “While some felt the pinnacle of recording technique had been reached, emerging technologies, notably multitrack recording, expanded the possibilities of what could be achieved in the studio by clever engineering and creative overdubbing. By 1966, popular recording had rejected any notion of fidelity to live performances in favor of studio creations, or what one producer called ‘the sound that never was’ (Schmidt, 2012).

From this point, the next generation of technology was developed, digital audio technology. It was first commercialized in the 70s, and it borrowed the technology of Pulse Code Modulation (PCM)[4] from the telecommunications field, employing it to record into digital audio workstations (DAWs). Later on, other inventions, such as the MIDI protocol (1982), which allowed for communication between synthesizers and sequencers and the fabrication of midi instruments, allowed artists to manipulate any sound from a keyboard, widening the creative possibilities and making musical creation more accessible (Cook et al., 2009).

Altogether, we can see technological advances not only changed the sound but turned the recording studio and the sound engineer into key figures in the creative process. “The recording studio became less like a photographer’s studio – everything carefully arranged to be captured in a sitting – and more like an artist’s studio, the finished product emerging from scraps and overdubs. To make a record was to make something quite new; no longer to pursue ‘fidelity’ to an original ‘live’ event” (Frith, 2001).

---

[4] PCM: Method developed at the end of the XIXth century used to digitally represent an analog signal. It is the method with which analog audio signals are transformed to achieve digital audio.

Ultimately came the internet, during the 90s, and music slowly became an easily accessible entertainment. With the internet, the means of consuming music changed: “By the late 1990s, the appearance of various file compression techniques and increases in the overall speed of computer networking allowed the digital distribution of music (both legally and illegally obtained) to become a mass phenomenon.” (Théberge, 2001).

These developments have culminated in consumption of music mainly through streaming platforms, online sites and social media. This, in turn, has an impact on the creative process, by pressuring artists to make catchy, short songs that capture the attention from the very first seconds[5]. This can be seen as a decline in the artistry of pop music or, oppositely, an opportunity for new ear-catching techniques and ideas to be discovered.

2.3 Sonic characteristics of Pop

With the arrival of the internet, it has become of key importance for music distribution companies to make their products be well positioned and visible in the vastness of the web. In consequence, today’s biggest music companies focus on promoting music under the category of pop, to be able to market it in pop charts, pop culture sites, streaming platforms’ pop playlists, etc. and reach as large an audience as possible. This ends up causing ambiguity in what the term pop means in musical terms, and a need to define its characteristics.

In 2003 the idea of Hit Song Science, a developing technology that explored the possibility to predict whether a song would be a hit by using machine learning, was first attempted by a Catalan company called Polyphonic HMI. The company founders eventually changed, but the idea was spread around the globe and, in 2011, scientists from the Intelligent Systems Laboratory in the university of Bristol, looked at top 40 charts of the last 50 years in the United Kingdom and classified them using multiple musical feature categories, such as tempo, time signature, harmonic simplicity, loudness, harmonic sequences, etc. The objective was to build a system that could discern which songs would be on the top 5 of the charts and which ones would be less popular (position

[5] Studies over Spotify’s skip data show that, the likelihood of skipping a song during the first 30 seconds is 35% and among teenagers it is a 50% with a rising tendency. Lamere, P. (2014). The Skip. 22/03/2022, from Music Machinery. Url: https://musicmachinery.com/2014/05/02/the-skip/

30 to 40). The accuracy of this system was found to be of a 60% chance of success, but it was also acknowledged that the algorithm should be periodically updated, in order to be in line with the music over time and continue to be accurate, because of the fast-changing nature of popular music. However, they found interesting facts, such as that, until the 90s, harmonic progressions of hits were simpler than non-hits or that, contrarily, time signatures in hit songs after the early 90s became simpler and binary (4/4 was established as the most common time signature). They also noticed that, with the passing of time, all songs in the chart became louder and that hits were relatively louder than the songs that had not been so popular. [6]

A similar study was carried out (Serrà et al, 2012), where their dataset was composed of nearly 500.000 songs spanning from 1955 to 2010 in a variety of popular genres. Along with the dataset, there were descriptors of each song covering loudness, pitch and timbre characteristics. The results they arrived to, surprisingly indicated that there had been no clear evolution in contemporary western popular music in more than fifty years. They did find a trend in songs becoming louder and potentially losing dynamics, as well as a homogenization of the timbral palette.

Some people argue that the drastic commercialization of the genre has affected the musical quality and they might have a point, looking at these studies. “Record companies and others who operate in it try as far as possible to reduce its uncertainties and their own risk - increasingly by promoting as “popular” the sort of music which will have the widest possible appeal..[] Thus, just as in the early days of radio, the resulting “popular” music is often an adulterated form of some more vital style, or, increasingly in recent times, is manufactured in the studio with a mass audience in mind.” (Martin, 1995). This tendency has induced, in part of the population, the thought that pop is a low-value genre, with the only aim to be economically successful. What is generally not taken into account in this type of opinion is the increased value of production and post-production of this music in recent years. Songwriting techniques have not fundamentally changed since the establishment of the genre, but production and post-production techniques have been

---

[6] Ni, Y. Santos-Rodríguez, R. McVicar, M. & De Bie, T. (2011). *Hit Song Science Once Again a Science?* 2022 March 22, University of Bristol. Url: http://www.tijldebie.net/publications/scoreahit_paper

exponentially developed and improved in the last years, thanks to advances in technology, allowing total sound manipulation. In general, since the times of The Beatles or even earlier, some pop music (though not an overwhelming amount of it) has challenged aesthetic, musical, cultural and technological ideas and practices.

In one of his videos[7], composer and youtuber David Bruce, talks about how the approach and mindset of modern pop producers has changed in the last few years, giving examples of big producers such as Ian Kirkpatrick (Dua Lipa, Selena Gomez,..), Noah 40 Shebib (Drake, Beyonce, Alicia Keys,..) and Finneas (Billie Eilish, Camila Cabello, Tove Lo,..), who approach producing almost from an engineer's point of view, in terms of spectrum instead of pitches, deciding instrumentation in terms of space and sonic density, caring to leave spectral space for the main elements, etc. Ian Kirkpatrick, for example, appears in a fragment of the video talking about developing a song's structure by saving up the highest range of the spectrum for the chorus, in order to create a splash of brightness and make the song unexpectedly bigger when that part arrives or removing melodic lines that could clash with the high end spectrum of the main vocal.
Bruce attributes this producer's attitude to the endless possibilities of the latest sound manipulation technology. He says it's like "playing with sound as a flexible lego", making his point that, in the end, the flexibility technology brings us allows for the use of full creativity, to go beyond typical sonic limits and toward the most surprising production, which is, in great measure, where the artistic value of today's pop lies.

[7] Bruce, D. [David Bruce Composer]. (2020, May 18). *The Subtle Art of Modern Pop Production*. [Video] Youtube. https://www.youtube.com/watch?v=dwexkS4Djb8

2.4 The production of music: roles and responsibilities

In the process of creating a song, multiple creative and technical figures are needed. The producer, the recording engineer, the mixing engineer and the mastering engineer are the people in charge of the technical aspects of musical creation.[8]

The producer is the figure in between songwriting and post-production. Once the song is written by the artist and/or their songwriter/s, the producer is hired to arrange the song. Arranging takes into account the selection of the instruments, their timbre, where they are placed (stereo, depth, virtual space, etc.), how they are played, and other very important creative decisions. The producer's job is to shape the structure and main sonic characteristics of the song, therefore, he/she must have both musical and technical knowledge, to create, shape and record a definitive demo of the song. It is not uncommon that a producer has an influence in the songwriting or mixing of a song, the roles are very closely related and the line can sometimes be blurred, but, in the end, it comes down to realizing the artist's vision.

After the song is produced it has to be recorded. Nowadays it is not unusual to find different elements of a song have been recorded in different studios. This is due to the popularity of multitrack recording in this genre (see Chapter 2.2), with which the sound is not affected as much by the acoustics of the space it is recorded in, (the instruments are practically isolated with close-miking techniques) and this allows for more post-recording flexibility.

After the recording, the sonic characteristics of the music are shaped by the mixing engineer. The mix engineer works from audio files (the recordings of the instruments + usually a rough mix made by the producer to show the artistic direction) after they have been produced and are definitive. A mixing engineer's job is to take the song files and process them (analogically or digitally), balance the tracks with each other, emphasize the importance of certain elements, make every instrument contribute to the song as a whole or look for contrast, manage transitions, etc. The mixer must match the sound the artist and the producer are looking for. Mixing is not a step by step process that is always

[8] Leviatan, Y. (2022). Making Music: The 6 Stages of Music Production. 2022 March 22, de Waves Audio Sitio web: https://www.waves.com/six-stages-of-music-production

done the same way, it depends entirely on the nature of the sounds that arrive in the first place and the desire for where to take them, as well as the taste of the mixer. “A mix should optimize the song, the vocal, the performances, the arrangement, and the engineering. It should sound good on a wide range of high- and low-end systems, and at any volume. It should be appropriate for the applicable format in a radio or club and hold its own in terms of apparent loudness or excitement with comparable material “(Burgess, 2013). The mixing process requires perceptive, technical and creative skills and it is of key importance in highlighting the emotion, movement and feeling of a song.
The processes that a mixing engineer carries out have to do with volume, timbre, dynamics, loudness and stereo image of the individual audio files and of the song as a whole, as well as adding creative effects when needed. Ultimately his job is to transform a set of separate audio files (instruments) into a balanced professional sounding song. Some typical tools used by a mixing engineer are volume faders, panning knobs, equalizers, compressors, harmonic exciters, limiters, FXs like reverb, delay, modulation, distortion, etc.

Once the track is mixed, the file (a stereo track, not a bunch of audio files) is sent to the mastering engineer, who furtherly shapes the final spectral and volume/dynamic details to get the song finished and ready for different reproduction means.

2.5 Common practices in pop mixing

Since the popularization of multitrack recording, the sound of popular music productions changed. “The process of multitrack recording has become the primary mode of production in popular music. It has resulted in both a particular ‘sound’ – dynamically compressed and spatially separated” (Schlemm 1982). Subsequently, mixing techniques changed as well. Some of the most elemental mixing standard techniques are related to space and volume.
In the year 2010, musicologist Allan Moore and Dr. Ruth Dockwray wrote in Popular Music magazine, the article “Configuring the sound-box 1965-1972” where they create the term sound-box, meaning the virtual space where popular music songs are played. Derived from this, they present 4 different typologies of mixes according to the positioning of the elements in space. One of these typologies, which they call Triangular, is one of the most basic mixing techniques used in today’s music. This is, placing vocals,

bass and drums in the center, highlighting them as the most important elements (melody and rhythm section), for they are key to understanding the song in different means of reproduction such as mono phone speakers, a club, a car or even the supermarket. Then, every other instrument is placed around this setting.

Another mixing technique, popularized in the XXth century, had to do with loudness perception. In the 1940s, engineers started to increase the volume of songs in order to make them stand out from others, to get the listeners attention in public spaces where radio and compilation records were being played. This practice developed into a phenomenon that was later named the loudness war. A loudness increase without clipping the audio signal is achieved by using compression and/or limiting. These tools reduce the dynamic range[9], resulting in dynamicless music and loss of contrasts and depth. The loudness war spread along more genres than pop, but typically, because pop is the most played genre in radio, clubs, stores, etc. and therefore, it is the genre that has incorporated it the most, for the need to make the songs continuously listenable in these noisy conditions.

Nowadays, streaming platforms have started to normalize[10] songs, in order to equal their levels and prevent the listener from having to constantly change the volume knob. These two examples show common mixing techniques in popular music, but, are there other standard procedures when mixing contemporary pop?

In a video[11] from the company Waves Audio, we can see Grammy-winning producer and engineer Scott Jacoby mixing a song and taking two different approaches, a pop-oriented mix and a rock type of mix. He goes through all groups of instruments and briefly explains his processing. For the pop mix he approaches a drum sound that is dry, deleting most of the room mics and deleting, as well, the overhead mics. He manipulates the envelope of the snare and kick drum, adding attack and reducing sustain, in order to hide room reflections and achieve a dry sound with an exaggerated transient. He processes the hi-

---

[9] Difference between the loudest and quietest moments of a song.

[10] Increasing/lowering the volume of a song to a target level. The relative dynamics of the song remain unchanged. For example, the normalization target level of Youtube is -14 LUFS, which means a song measuring -16 LUFS will be increased by 2 LUFS and one measuring 12LUFS will be lowered 2 LUFS.

[11] Jacoby, S. [Waves Audio]. (2019, October 21). 1 Song, 2 Mixes. How Mixing Can Transform a Song. [Vídeo] Youtube: https://www.youtube.com/watch?v=RJX-V1F8j0w&t=204s

hat the same way and adds distortion and high-end equalization to make it brighter. The words he uses to describe the pop mixed drums are tight, punchy, contained and sparkly. Contrarily, in the rock mix, the drums are made to sound roomy and big, using parallel compression. For the pop bass, he adds a high-pass filter to remove the high end and most of the mids, and adds harmonics in the sub-low and low section of the spectrum, leaving space in the low-mids and mids for other instruments, while achieving a very focused and boomy low end. In the rock mix, the bass fills a lot more of the spectrum, it is more present and defined. The pop acoustic guitars are made very bright and grainy with a tape emulation plugin and the electric guitar brings depth to the mix and it is used as a textural element, that sits behind the vocals to try to avoid masking of common frequencies, and to make the vocal seem more shiny and upfront, whereas, in the rock mix, it is more present and central. The processing of the vocals is not given a differentiated treatment in each mix, but the master bus processing is. The pop mix is equalized using a "smiley face" shape, adding low and high end, and it is also more compressed and limited.

The rock mix is maintained more natural and compressed with vintage emulations of analog gear, giving it a different timbral character.

We can find plenty of tutorials and audio education platforms where professional mixers talk about techniques for mixing pop and probably the general idea we can get from them (knowing that every mix is unique and that no rules are applied always in the same manner) is that drums must be dry, focused on the transient and impact, not on replicating a real "drum kit in a room" sound. The overall sound of the kit is not generally as important as its individual parts. The bass must be controlled and fill the low end while leaving space for everything else. The rest of the instruments must be filling the frequency and stereo spectrum, without overshadowing the vocals, which must be shiny and on top of everything. The overall mix is compressed so that energy is kept high during the whole track.

3. Methodology

Oftentimes the matter of music analysis is instantly related to the musicology field. The musicological approach to analysis consists in studying how the music is constructed from musical concepts such as songwriting, harmony, score, lyrics, performance, etc. and their relationships. Instead, here I attempt a sonological approach. Sonic analysis seeks to describe and explain the decisions and technological processes present in the recording and post-processing of music, which shape and transform it, adding an additional layer of meaning into the music.

In the sonological approach to musical analysis, the term “critical listening” is very present, as the ability to identify, by listening, the characteristics of sounds in recorded music. The existing methodology on critical listening, such as the aforementioned “Recording Analysis: How the record shapes the song” (Moylan, 2020) or others, such as “Critical Listening Skills for the Audio Professional” (Alton, F. 2006) or “Audio Production and Critical Listening” (Corey, J. 2016), are very comprehensive sources of information to take on the task. There are, as well, several websites, such as Soundgym[12], or Technical Ear Trainer[13] that help train the critical listening skill.

I consider, that in order to take on critical listening, knowledge on sound and its characteristics is essential. Therefore, the aim of this methodology is to distill the elemental information found on the references and develop a simplified methodology, providing the reader of a structured process to describe the fundamental aspects of each sound and its processing. The goal is to help distinguish individual sound sources and derive their relationships with others, to gain a profound understanding of the sonority of a piece of music and develop the ability to describe it. The focus of the methodology lies in the analysis of the recordings and their post-processing, although attention to songwriting and production is also present.

I begin the methodology by presenting awareness on listening biases and environment conditions, in order to get as close as possible to an ideal critical listening situation. Later, I introduce the matter of vocabulary for description, giving sources from where to obtain

[12] https://www.soundgym.co/
[13] https://webtet.net/apcl/#/

accurate definitions to use when describing timbrical, dynamic and spatial characteristics of instruments. I continue with the presentation of the sonic elements to be analyzed and the tools that can aid our listening with visual and numeric information, as well as presenting different collected graphical representation methods from existing books and a personal representation proposal. Lastly, I display the structure that I will use in the practical analysis cases.

The method is centered on analyzing contemporary pop music. Many of the processes carried out can be extrapolated and adapted to the analysis of other genres, it is not the tools used which must change, but the approach and focus of the analysis. The focus here lays on the vocal and rhythmic elements, for, in them, is where the essence of pop music lays. A catchy melody and a danceable beat are the DNA of any successful today pop songs. Because of this, the descriptions of timbre, space and loudness are made by separating the music in three main groups: vocals, rhythm and music, encompassing, on the latter, all the melodic and harmonic elements that support the main melody, giving them more relevance if they take a leading role.

3.1 Biases acknowledgments and suitable conditions for critical listening

As humans, subjectivity is inherent in our attempt to analyze and understand the world. Separating education, personal opinion, cultural baggage, surroundings, etc. from the experience of the world is unfeasible and in order to analyze any material, acknowledgement of biases is crucial. The writer of an analysis has the responsibility of defining a methodology that intends to diminish biases, where possible, and acknowledge and present them during the analysis. At the same time, the reader should adjust their environment and raise awareness of their own existing biases, as well as the writer's, in order to fully engage with the study.
In the book "Recording Analysis: how the record shapes the song", Moylan presents two types of biases we can imprint on a critical listening analysis: Personal biases and playback biases. The personal biases are conformed by our experience and relationship to the music, for example, our relationship to genre, previous information on the artist's work, previous listening of the song, past experiences related to the song, etc. as well as our level of education on critical listening. Through education we can learn to perceive

elements we would not have noticed previously, and, in this way, enlarge our musical experience. The second type of bias is listening conditions. The gear, space and position from which we analyze music are key to the ideas we derive from it, and therefore have great power to cloud our conclusions.

"We would not study a painting—a masterwork by, say, Michelangelo or Monet, Degas or Dali—by viewing it through a veil, or from position distinctly closer to one side than the other. We would not place it in a shadow-streaked location, or in poor lighting that does not reveal nuance in its broadest shapes." (Moylan, 2020)

As we would with a painting, we must find the most unbiased environment for critical listening (always knowing that the ideal listening situation doesn't exist).With the purpose of recognizing these biases and adjusting the listening conditions, here I present the listening conditions I utilized in my analyses, extracted from William Moylan's method and adjusted to my own environment.

- Level: Moylan advises listening at 82-85dB SPL with peaks up to 90dB SPL, "the because of the nonlinear relationship between frequency and loudness, as depicted in the Fletcher-Munson curves[14]. After testing this with a sonometer, I consider it is too loud for long exposures of time and it results in ear fatigue. Therefore, I adjusted the level to 70-75dBs SPL (with peak level up to 80 dBs). This level must be measured at listening position with a sonometer (in this case a PHONIC PAA3) set to A-weighting curve[15]. Floor noise of the room is <30dBs SPL.
- Room & position:
    - Studio: 5x7m, professional acoustically treated room. Listening position 2m away from the monitors.
- Home: 2.1x3x2.3m not acoustically treated room. Listening position <1m away from the monitors.
- Gear:

---

[14] Fletcher Munson Curves, also known as Equal Loudness Curves or Isophonic Curves, are the result of the study of human loudness perception. These curves show how humans have different loudness perception depending on pitch (frequency), by showing the amount of volume (dB SPL) needed between two frequency bands in order to perceive them equally loud.

[15] Loudness measurement curves are made to accurately measure volume according to human perception, given that we don't perceive the same loudness level for every frequency. The weighting curve must be adapted to the listening level.

  - Studio: Genelec 8040A active monitors (90W Bass + 90W treble); 45Hz -21kHz (-6dB) frequency response. 19mm tweeter, 165mm woofer, Genelec 7040A subwoofer (30Hz - 90 Hz (-6dB)).
  - Home: JBL 305 mkII active monitors (41W tweeter + 41W woofer); 49Hz - 20KHz frequency response. 25,4mm tweeter, 127mm woofer.
  - Headphone reference: Audio Technica M40X closed back.
- Media format: FLAC, lossless digital audio format at 16 bit 44.1kHz.

According to Moylan, headphone analysis is not advised due to:

- Balance is shifted as the amplitude doesn't get reinforced by the opposite ear (interaural level differences)[16].
- Timbre is shifted as frequency response of headphones can't reproduce low frequencies the same way as loudspeakers, because they don't' move big masses of air[17] and the "vibrational" aspect of low frequencies affecting the whole body is also lost.
- The suppression of one's Head Related Transfer Functions[18] (HRTFS) and room reflections affect our perception of space, stereo image, distance positions, timbre details

Given all these reasons, we shouldn't rely on headphones for a whole analysis, although they can be used to perceive small details like timbre changes, performance details, individual instruments' characteristics, etc. or for a change of perspective.

---

[16] There are some technological approximations to simulating, with headphones, the interaural cross-talk that occurs when listening with speakers, although they are not fully accurate because they do not take into account the effects of the physiognomy of each person.

[17] Physically, low frequencies have long wavelengths. The wavelength is the distance between two corresponding points in a periodic wave. Wavelength is defined by: $\lambda=v/f$, v being velocity of air, and f being frequency. If we calculate this, the wavelength of 40Hz would be: $\lambda=343/40\sim8,6m$. Thus, low frequencies in headphones can't be accurately represented because of the lack of air between the headphone membrane and the ear.

[18] Head Related Transfer Functions are the unique responses, dependent on body physiognomy (ears, nose, cheekbones, shoulders, etc.) that affect the frequency, amplitude and time domain of incoming sound waves that hit our body. This phenomenon shapes sound in a particular way for each individual before it enters the ear canal. As we grow, we instinctively learn to hear the world through our HRTF filter.

3.2 Common vocabulary for sound description

For the purpose of sharing knowledge on sound, a physical event, we must agree on a terminology to translate perception into words, with the purpose of sharing our ideas, sensations and opinions. There are commonly used words and phrases generally accepted by the community of musicians and engineers, although confusions aren't uncommon and we can often find people using the same terms to describe different ideas, or contrarily, a lack of vocabulary knowledge from people who want to express their sound-related ideas and don't have the appropriate language to transmit them. As Moylan summarizes:
"our custom is to describe sound by using analogy and by using cross-modal terminology from other senses. We resort to words such as 'warm,' 'round,' 'fat,' 'mellow,' 'percussive,' 'edgy,' and a great many others—imprecise at best, and typically grossly inadequate and ineffective; often misleading, and commonly merely meaningless jargon." (2020, Moylan).

In order to standardize sensory audio vocabulary, there have been several attempts to construct a lexicon for sound description. The most validated one being a report from March 2017 by the International Communication Union, titled "Methods for selecting and describing attributes and terms, in the preparation of subjective tests" (report number: ITU-R BS 2399-0).

For this report, an Audio Wheel[19] was conformed (see Fig. 3.2.1) as well as a table of sound attribute definitions, which grouped into 6 categories (loudness, timbre, dynamics, spatial, transparency and artefacts) the elemental words for describing a sound.

[19] "A sound wheel is a hierarchical visual representation of a perceptual attribute lexicon. Attributes with similarities are placed in the same category. Similar categories are grouped near each other and placed on a doughnut graph. The result is a visual representation of how attributes are related. The wheel may be used to explain and understand descriptive sensory analysis and it may be used for panelists during training." ITU-R. (March 2017). *Methods for selecting and describing attributes and terms, in the preparation of subjective tests*. Report. Geneva, Switzerland: International Telecommunication Union Radiocommunication Assembly (ITU-R).

Figure 3.2.1: Audio Wheel. Audio terminology circle-graph collected. Source: ITU- R BS 2399-

Some of the most commonly misunderstood or misused terms that I have found in my musical journey are in this report described as:

- Punch: whether the strokes on drums and bass are reproduced with clout, almost as if you can feel the blow. The ability to effortlessly handle large volume excursions without compression.
- Powerful: The ability to handle high sound levels, especially when striking the drums and bass. Indicates whether the Punch, Attack and Bass precision are maintained at high volume.
- Brilliance: Treble or high frequency extension: A little: As if you hear music through a door, muffled, blurred or dull. – A lot: Crystal-clear reproduction extended treble range with airy and open treble. Lightness, purity and clarity with space for instruments. Clarity in the upper frequencies without being sharp or shrill and without distortion.
- Nasal: A closed sound with pronounced midrange. Gives the impression corresponding to vocalists singing through the nose (nasal).

- Boomy: Resonances in the low bass, as sound in a large barrel, which gives a prominent bass resound resounding (reverberating) when bass and bass drums are heard. The representation tends to become muddy and imprecise.
- Full: If both low and high frequencies are well represented with good extension the sound is Full.
- Depth: The radial extent of source, scene or ensemble from the listener (in any direction away from the listener).
- Clarity: The impression of how clearly different elements in a scene can be spatially distinguished from each other.
- Presence: Does it sound as if the sound sources are present and not distant or absent?
- Clean: It is easy to listen into the music, which is timbral clear and distinct. The opposite of clean: dull, muddy.
- Rough: A hoarse off-sound unintentionally accompanying the reproduced sound. Bass distortion.

These terms are used in my analyses, in chapter 4, as well as others from other valuable fonts such as:

- Sound on Sound vocabulary glossary: Includes sound description vocabulary as well as audio engineering technical terms. It is regularly updated and welcomes email suggestions. https://www.soundonsound.com/glossary
- The Semantic Space of Sounds, Torben Pedersen (2008): Includes a whole lexicon of descriptors and attributes of sound. Each word is scaled between 0-10 points regarding different categories. Url:https://www.researchgate.net/publication/263964081_The_Semantic_Space_of_Sounds
- The Art of Mixing, David Gibson (1997): Mixing guide that includes charts of vocabulary used when equalizing, relating them to specific instruments' timbres (Fig. 3.2.2).
- The Relentless Pursuit of Tone, Fink, R. Latour, M. & Wallmark, Z: Includes vocabulary and graphs related to the frequency spectrum (Fig. 3.2.3).

Frequency	40-100	100-200	200-800	800-1000	1000-5000	5000-8000	8000-12,000
SOUNDS							
BASS	Bottom	Roundness	Muddiness	Body on Small Speakers	Presence	High End	Hiss
KICK	Bottom	Roundness	Muddiness			High End	Hiss
SNARE	X	Fullness	Muddiness			Presence	X
TOMS		Fullness	Muddiness		Presence Irritation	High End	X
FLOOR TOMS	Bottom	Fullness	Muddiness		Presence		X
HI-HAT CYMBALS	X		Muddiness Bleed		Irritation	Clarity/ Crispness	Shimmer/ Sizzle
VOICE	Rumble	Fullness	Muddiness		Presence Irritation Telephone	Clarity/ Crispness Sibilance-6K	Sparkle/ Hiss
PIANO	Bottom	Fullness	Muddiness	Muddiness	Presence	Clarity/	Harmonics
HARP		Pedal Noise			Twanginess	Crispness	
ELECTRIC GUITAR	X	Fullness Crunch	Muddiness Roundness		Cut/Shred Irritation	Crispness Thinness	Hiss
ACOUSTIC GUITAR	X	Fullness	Muddiness			Clarity/ Crispness	Sparkle
ORGAN	Bottom	Fullness	Muddiness			Clarity/ Crispness	
STRINGS	Bottom	Fullness	Muddiness		Irritation Digital Sound	Clarity/ Crispness	Sparkle
HORNS	X	Fullness	Muddiness	Roundness		Clarity/ Crispness	
CONGA	Boominess	Fullness				Clarity/ Crispness	
HARMONICA	X	Fullness			Irritation	Clarity/ Crispness	X

Chart 4. Equalization Chart

Figure 3.2.2: Equalization chart. Shows terminology related to different instrument and their frequency spectrum. Source: Gibson, D. (1997).

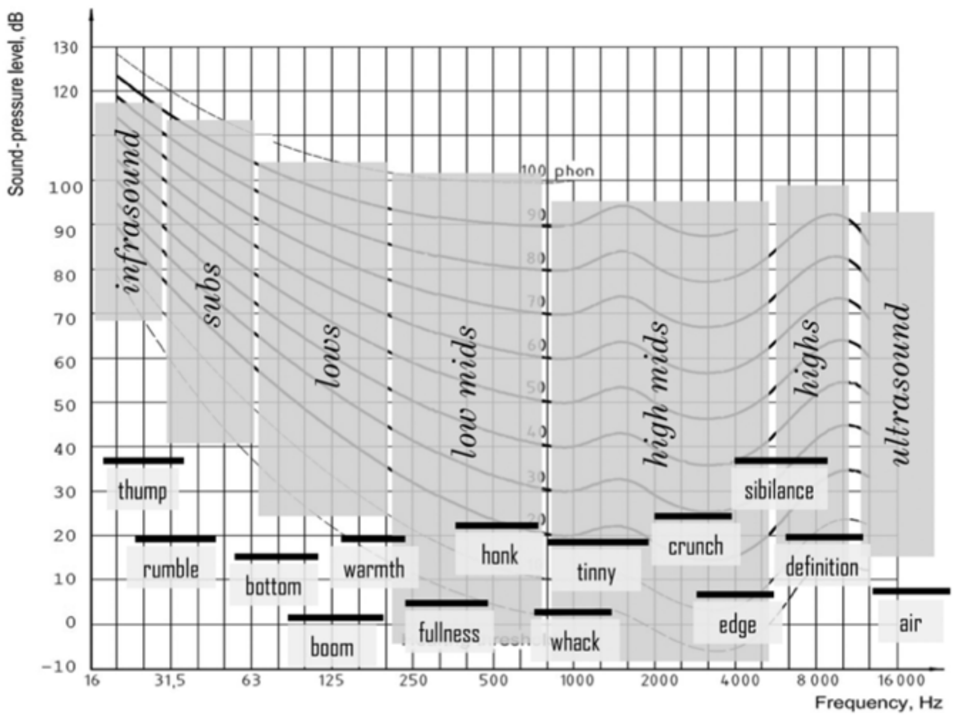

Figure 3.2.3: Divisons of the audible spectrum. Shows terminology for frequency bands. Source: Fink, R (et al. 2018).

3.3 Analyzable musical aspects

With the purpose of making the analysis comprehensible we must limit the elements under analysis, create a framework of key elements to explore and in that way, obtain the necessary information to perform a full analysis. In his book, William Moylan decomposes the analyzable aspects of recorded music in 3 main categories: timbre, space and loudness. Each of them has intrinsic qualities and tools to visualize them, that are presented below.

Timbre:

Timbre is a conjunction of frequency and loudness characteristics that together form an auditory sensation which allows us to distinguish and recognize sounds, for example, discerning different people's voices or different instruments. Timbre is a quality that changes over time and it is composed of 3 main elements:

- Spectral content: Energy distribution of the partials[20] that conform the spectrum of the sound.
- Dynamic envelope: Amplitude variation of the sound over time. The dynamics envelope is usually called the ADSR envelope because it is composed of 4 parts: attack, decay, sustain and release (Fig. 3.3.1).
- Spectral envelope: Amplitude variations of the partials over time.

Although the latter one is found to be less significant:

"Research that uses multidimensional scaling (MDS) algorithms to discover the acoustic correlates of listeners' perceptions timbre suggests that the most perceptually salient acoustic components of timbre are a combination of spectral center of gravity (the dispersal and strength of frequencies above the fundamental frequency) and attack time (the nature of the onset of a sound)[...] The change in dispersal and strength of harmonics over time may contribute to listener perceptions, but this has been found to be less significant than the aforementioned elements" (McAdams, 1999).

[20] Partials: any of the simple tones that compose a complex sound. The lowest partial of a sound is called the fundamental, and the partials above it are multiple frequencies of the fundamental, which conform what we know as pitch. When the partials are exact multiples of the fundamental, we call them harmonics. Some instruments are conformed of only harmonic or inharmonic partials and some are conformed by a combination of both.

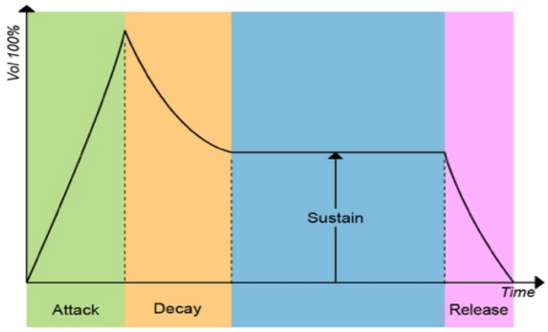

Figure 3.3.1: Dynamic envelope of a sound (ADSR Curve). Source: Cómo crear sonidos propios utilizando la envolvente adsr. Djp Music School. Madrid. url: https://djpmusicschool.com/2020/11/como-crear-sonidos-propios-envolvente-adsr/

Tools for timbral analysis:

In order to see the spectral content of an instrument or group of sounds, a mathematical process called a Fast Fourier Transform (FFT) needs to be carried out. This process splits a signal into its conforming frequency bands. The tools we can use to graphically display an FFT are:

- Spectrum analyzer: It shows an x-y graph that displays frequency on the horizontal axis, amplitude on the vertical axis and changes shape during time (usually without the possibility to store data in time). It is useful for short passages of music.
- Spectrogram: It shows an x-y graph that displays time on the horizontal axis, frequency on the vertical axis and amplitude shown by color (usually from blue (low amplitude) to red (high amplitude)). It is more detailed, which makes it good to observe performance details and it is best for analysis over time.

These tools can and must be configured depending on the aspect we intend to see:

Configuration of spectrum analysis tools:

- Window Shape: "Windowing consists of multiplying the time record by a finite-length window with an amplitude that varies smoothly and gradually toward zero at the edges. This makes the endpoints of the waveform meet and, therefore, results in a continuous waveform without sharp transitions”[21].

The plugin Voxengo Span we can choose between Blackmann (Fig. 3.3.2) or Hann (Fig. 3.3.3) window. The Hann window is narrower and has higher lobes which may be good for transient visualization while the Blackmann is wider and better for more sustained sounds.

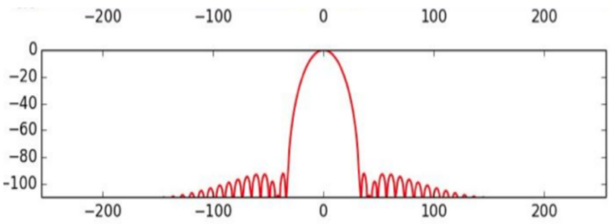

Figure 3.3.2: Blackmann window shape. Source: Barrachina, A. Anàlisi síntesi I processament del so 2: Apunts SFTF.

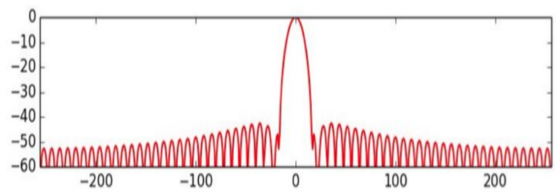

Figure 3.3.3: Hann window shape. Source: Barrachina, A. Anàlisi síntesi I processament del so 2: Apunts SFTF.

- Window Size: The window size is the number of samples considered in every window frame. For a general view of a song's spectrum we can use small sizes such as 1024 or 2048. For low frequency resolution we need to use big window sizes (8192 and up). For transient detection (such as drums) we should find a middle ground that doesn’t compromise the temporal aspect but allows us to see low end information (e.g. if there’s a kick drum). A larger overlap between analysis frames (or window averaging, as I mention below) is usually a good trick here.
- Type of averaging: We can average the spectrum in time, in order to see a somewhat-constant shape and get an idea of what a section’s timbre is like. For upbeat pop music, a good measure of spectrum averaging could be between 400 and 500ms, although it will depend on the source.

[21] National Instruments Corp. (2022). *Understanding FFTs and Windowing*. April 18 2022, National Instruments Corp. Url: https://download.ni.com/evaluation/pxi/Understanding%20FFTs%20and%20Windowing.pdf

- Range of display: The range is the minimum and maximum frequency for the analyzer to take into account. Frequency range is dependent on Sample Rate frequency[22], due to Nyquist's limit[23].

We need to be aware that aural and visual perception are not aligned. Due to physical phenomena (as seen in Chapter 3.1) sound level is not equally perceived in all frequency bands. Therefore, what we see in a spectrum analyzer will never truly coincide with our sonic perception and we must modify the settings to best translate every input signal into reliable graphs.[24]

Space:

The sense of space we receive from our auditory sense is defined by:

- Psychoacoustic phenomena:
  - Interaural Time Differences and Interaural Level Differences: An audio source arrives at different times and volumes to each ear, due to the distance between them and the acoustic absorption of the head. Our brain interprets these differences and is able to calculate the original audio source and place it in space.
  - Timbre changes related to distance loss: "As sounds move further from us, higher frequencies diminish more rapidly than lower frequencies— being absorbed by the air, attenuated by air friction. [] Timbre is fixed when the source is recorded. Raising the amplitude level of a source in the mix does not change distance, when the changing loudness is not accompanied by a change of low-amplitude detail in timbre—loudness changes without timbral detail changes will not shift distance location" (Moylan, 2022).

---

[22] Number of samples per second taken from a continuous signal

[23] Nyquist's Theorem states that in order to properly represent a waveform, a sample rate of at least twice the signal frequency is required. Signals above the Nyquist frequency (fs/2) are folded back to appear as lower frequency signals, causing wave cancellations known as aliasing. To avoid this, an anti-alias low-pass filter is used.

[24] In This IRCAM webpage, you can find accurate explanations of further FFT parameters: https://support.ircam.fr/docs/AudioSculpt/3.0/co/Window%20Size.html

- Reverb: Reflections from sound hitting the surfaces of an enclosed space. The time delay these reflections add and the timbral changes due to acoustic properties of the space[25] are summed to the direct sound, changing the perception we get from it.
- Echo: When the first reflection of a sound is delayed more than 35ms from the original initial wave, we perceive it as a different tone. This phenomenon is called "Haas effect".

The stereo image of recorded music utilizes these mechanisms to make the brain interpret space. We get the sense of something being left or right because the information provided by the left and right speakers is not the same. The center of the stereo (also usually called Mid) is interpreted by our brain from the information that is identical in both channels (we say it is correlated information), whereas, what appears to be on the sides or outside the speakers, is the difference (the uncorrelated information) between them (forming the image of what we call the Sides).

To describe space in recorded music we can focus on the overall sound, the dimensions and characteristics of the whole sound stage, as well as the distance location, lateral location and width of the individual sound sources. In this way, we can state whether a sound is far/close, wide/narrow, long/short, bright/dark, rich (very reflective)/dry, etc.

There is a way to measure correlation (in other words, measuring how wide the stereo image of a piece of music is), by using correlation meters (Fig. 3.3.4). When the meter is all the way to +1, the signal is completely correlated (mono) and when it displays -1, it is completely on the sides (nothing in mono). Every number in between is a combination of mid and sides, what we know as stereo. Another form of spatial representation are Lissajous figures. These figures are displayed by devices called Goniometers. In a goniometer (Fig. 3.3.5), when a signal is completely mono, a straight vertical line is drawn, contrarily, when a signal is totally uncorrelated (only has sides information), a horizontal line appears. Every figure in between is a combination of both.

[25] Absorption and diffraction are acoustic phenomena that occur when sound hits any surface. Depending on the surface, the amount of absorption or diffraction changes and they can happen both at the same time.

Figure 3.3.4: Brainworx balance and correlation meter. Source: Screenshot

Figure 3.3.5: Lissajous figure. Source: Screenshot of Steinberg's Gonio3 goniometer plugin

These attributes of sound may be used by sound engineers to manipulate the spatial information and create fake scenarios. In today's pop music, correlation is a very important matter, since listening to music on phones, computer and Bluetooth speakers (that have little to no stereo, because they usually have only one speaker or they have two very close speakers that can't portray a faithful stereo image) has become a generalized practice, mix engineers have had to make mono compatibility a priority, making sure that spectral information and timbre are not lost when listening in mono.

Since multitrack recording has become the standard method of recording pop, the dry and manageable sound of the recordings has allowed for experimentation on the creation of space. The sense of space can be recreative or creative. In pop music today, what we usually hear are manipulated spaces, carefully set to best suit each instrument, regardless of the "realness" of the space. It is not unusual to encounter several different spaces in the same song or even applied to the same instrument (i.e. different reverberation for each elements of a drum kit). Reverbs can also be processed with equalization, distortion or any creative processing, leading to soundscapes that would be impossible in real life.

Loudness/Dynamics:

Loudness is a measure related to the human perception of sound pressure level. Sound pressure level translates to the intensity of sound (volume) that hits our ear. Intensity of sound is measured in Pascals (Pascals are the universal system's unit to measure pressure, used on sound because waves are, essentially, pressure changes traveling through the air) but decibels of sound pressure level (dBs SPL) have been adopted as a standard unit for loudness measures, for practical mathematical reasons. The dB SPL scale ranges from the limits of human audition regarding level, where the maximum average level that can be endured before the threshold of pain is 20 Pascals = 120dBSPL and the lowest is 20 microPascals = 0dBSPL.

Dynamics are a comparison between the loudest and quietest parts of a piece of music. The amount of volume between the loudest and quietest part can be expressed in dBs. We usually encounter that the music with the most dynamic range is classical music, with up to 20 or 30dBs between the loudest and quietest passages. Instead, in pop music, the dynamic range is usually small, because the music has an energetic and danceable nature, and is usually compressed and limited to get this effect.

There are different units to express loudness and Dynamics values:

- dB FS: Measure of amplitude level in digital systems. 0dBFS are defined as the maximum level achievable before digital clipping.
- LUFS/LKFS: Loudness Unit relative to Full Scale or loudness K-weighted related to Full Scale. They are both the same measure of amplitude, linked to human perception using the K-weighted equal loudness curve (see Fig. 3.3.6) at a 70dBA volume. They have different names because they were defined by different organizations, EBU R128 and ITU-R BS.1770.0 1LKFS=1LUF
  - LU: Loudness units related to equal loudness curves, not related to Full Scale. It is not an absolute measure. It is used for comparison (usually comparison of dynamic range between songs).

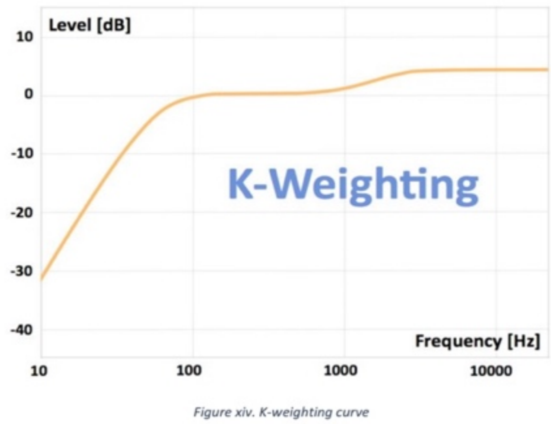

Figure 3.3.6: *K-weighting loudness curve.* Source: Camerer, F. (2020) Loudness – an Audio Levelling Revolution. EBU R 128 report

Using these units, we can express different loudness and dynamics measurements:
The dBFS untit does not take human perception into account, therefore the use of LUFS units is advised.

<u>Measures of loudness:</u>

- Peak and true peak value: Peak value refers to the loudest point of an audio signal. Peak value is measured in dB FS and using peak meters. In the last years, the peak value measure has been substituted by True Peak value, which measures the loudest point of a signal in the reconstruction process when the signal is converted to analog[26]. This is a more accurate measure than traditional peak metering because it makes sure that there's no clipping at any point of the DAC process. True peak meters use the dB True peak unit (dB TP).
- RMS: Average power of the audio signal. Measured in dBV or dBU[27].
- Integrated loudness/Program Loudness: Average loudness (related to human perception loudness curves) of a whole piece of audio. Measured in LUFS.
- Short-Term loudness: Average loudness of the last 3 seconds of audio. Useful for passages of audio. Measured in LUFS.

---

[26] The Digital to analog conversion process (DAC) is based on interpolating the various samples that define the waveform (the number of samples is determined by the sample rate of the file) in order to reconstruct the wave. True Peak detection was found necessary when realizing that a 0dBFS signal could still clip when reconstructed for playback.
[27] Units of voltage measurement.

- Momentary loudness: Average loudness of the last 400 milliseconds of audio. Measured in LUFS.

Measures of dynamics:

- Peak to Loudness ratio (PLR): Ratio between true peak value and the gated[28] integrated loudness value. "we measure the integrated loudness, but if that loudness gets too close to full scale, or in many cases too close to -1dBTP (TP stands for True Peak), then it means that probably a lot of compression or limiting has been applied and that's an indication that possibly you might start to lose transients and clarity." (Bob Katz, 2021).
- Peak to Short Term ratio (PSR): Ungated peak to Short Term Loudness Ratio measurement. Useful for transient information. "High relative PSR values suggest wide dynamics. Conversely low relative PSR values suggest reduced dynamics, excessive limiting, and elevated perceived loudness" (Shepherd et.al, 2017)
- Loudness Range (LRA): Gated measure of integrated loudness variations. (the top 5% and the lowest 10% of the total Loudness Range is being excluded from the measurement). "Loudness Range is supplementary to the main audio measure, Program Loudness, of EBU R 128. Loudness Range measures the variation of loudness on a macroscopic time-scale" (AES, 2022) It is intended to measure long passages, as the dynamic range of the average loudness. Measured in LU.

Sonic analysis should always take into account loudness measurements instead of RMS values, for they don't take psychoacoustics into account, and therefore, are not true to our experience of the music. For example, a low frequency signal could have a high RMS value while it would not be perceived as loud, because of the psychoacoustical phenomenon reflected in Fletcher & Munson curves.

---

[28] A measure that omits sudden increases in amplitude or periods of silence (when the signal is <-70 LUFS). (Audiokinetic Inc. (2022*). Loudness Measurement Scopes*. 22 April,2022.Url:https://www.audiokinetic.com/library/edge/?source=Help&id=loudness_measurement_scopes).

All of the loudness measures presented above can be obtained by using loudness meters such as Youlean Loudness Meter, Izotope RX Waveform Stats, Waves WLM, SPL Hawk Eye, etc.

3.4 Digital visualization tools

As we have seen, there are tools with which to aid the auditive analysis. It is clear that, in sonic analysis, a trained ear is the best tool we can have, for the experience of music is attached to perception. Many hours of ear training are required to be able to perceive the smallest nuances in music and there have been some very useful ear training platforms developed for sound engineers, such as Soundgym[29], Technical Ear Trainer[30], or Match the Mix[31] by Sonarworks, which can help us develop our listening skills. We can also find more in-depth information in books such as "Audio Production and Critical Listening: Technical Ear Training" (Corey & Benson, 2016) or previously mentioned "Recording Analysis: How the Record Shapes the Song" (Moylan, 2020).

It would be an ideal situation to be able to describe all details of a piece of music only with our listening skills, but this is a long-term goal. In the meantime, it can be fruitful and helpful for the learning process to complement our attentive listening with complementary digital tools, which can make us realize of things we might not have heard. We must know how these tools work, given that it is crucial to set up visualization tools correctly because they will condition our conclusions.

In my analysis (in Chapter 5) for the spectrum visualization I utilized Hann window for the percussion or transient-filled elements, and Blackmann window for the rest. The FFT size oscillates between 4096 and 16384, depending on the sound source and the level of detail that is pursued. Generally, I used higher resolution values to visualize low end information, whereas when analyzing the high end, lower values worked better to get an average view and not an excessive amount of small peaks.

[29] https://www.soundgym.co/
[30] https://webtet.net/apcl/#/
[31] https://matchthemix.sonarworks.com/

These are the tools that I use, grouped by the sound attribute they display[32]:

Timbre:

- Spectrogram:
    - Izotope RX8
    - Sonic Visualiser
- Spectral analyzer:
    - Voxengo SPAN
    - Izotope RX8 Average Spectrum
    - Fabfilter pro Q 3

Space:

- M-S listening: Brainworx Control
- Correlation meter: Brainworx Control
- Logic Multimetre's Goniometer or Izotope Insight Goniometer

Loudness/dynamics:

- Loudness measures:
    - Youlean Loudness Meter
    - Waves WMN
    - Izotope RX Waveform Statistics
    - Levels by Mastering the Mix

Overall tools

- Peel by Zplane: Display of stereo image related to frequency band

3.5 Graphical representation

With the aid of visuals, we form a clearer idea of how the sonic elements are placed and related. Several authors have created graphical representations of their analysis findings, and there are some very interesting approaches that join two or more attributes, which help understand the relationships between them or display connections, similarities or differences we wouldn't have found otherwise.

These graphs are presented in Appendix 1. I find Moylan's ("Recording Analysis: How the Record Shapes the Song) to be very complete, but the lack of color and the verticality, make the information hard to read, whereas Frampton's model ("How Pros Make Hits") presents a more minimalistic style, using color and intensity of color as an asset to bring information without crowding the page, although his graphs are not as exhaustive.

From these graphic representations I created a new model (Fig 3.5.1), with the intention of joining the ideas I find interesting in the above-mentioned references, achieving a representation method that contains various information and, at the same time, is easy to

[32] All links to the plugins' pages are on Appendix 2.

understand. This graph, merges the two approaches in regards to spatiality and frequency spectrum and additionally can make every section of the structure be read from left to right, as a book. Each instrument is given a color, for clear and direct relationships and the opacity of color is related to the depth of the instrument in the mix (more opaque= more present).

This example is not related to a specific song. Due to lack of time, here I present the fundamental idea, with the objective of developing this line of work in the future.

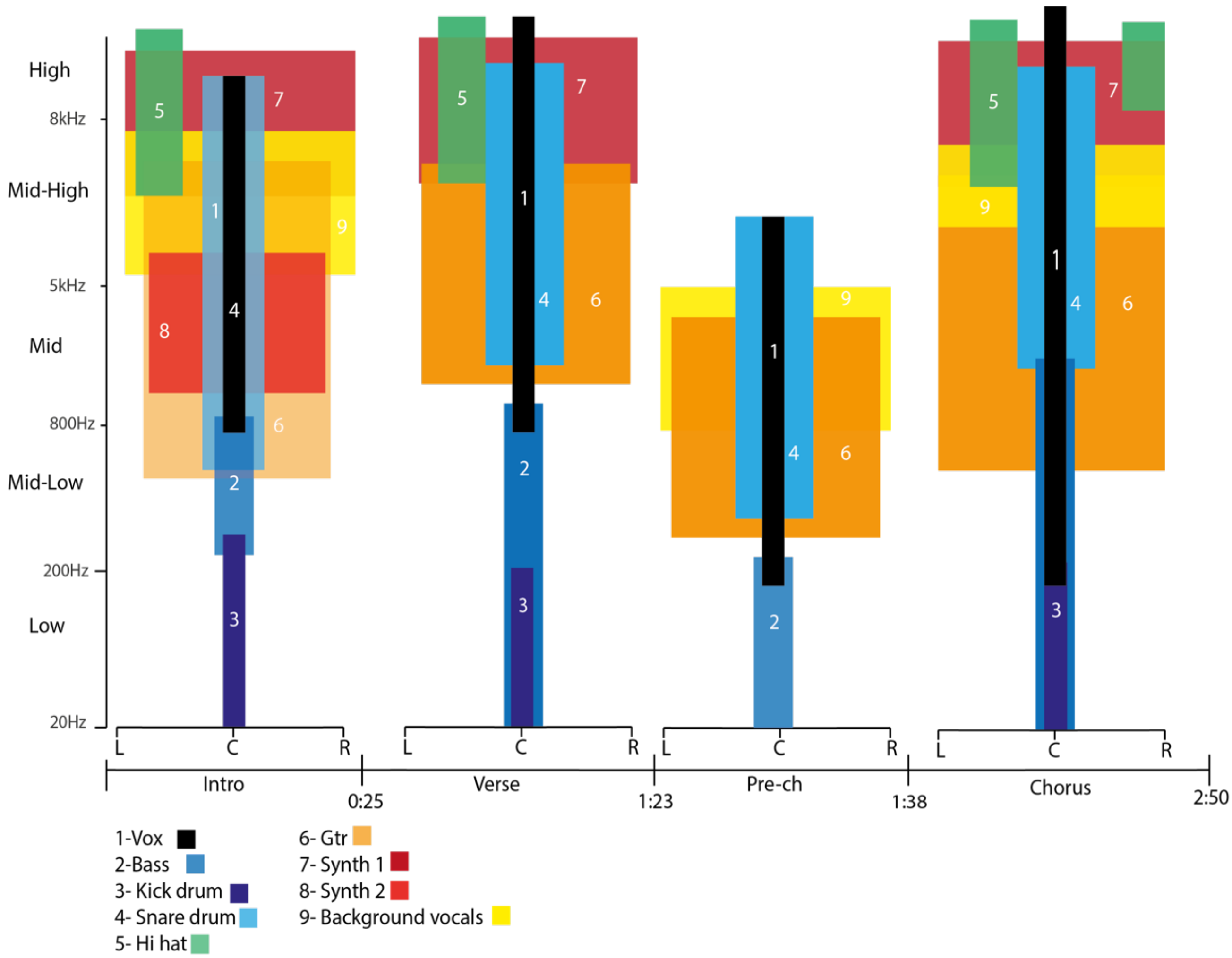

Figure 3.5.1: Graph of spectrum and stereo field relationship for a fictitious song. X axis: Stereo field and song structure Y axis: frequency spectrum Color: Instrument Opacity: depth of instrument in the mix.

3.6 Structure of analysis

Before the analysis, stems[33] of the track should be extracted. I extracted stems from AI-powered software LALAL.AI[34], a high-quality online stem separator. After comparing it with Izotope RX's music rebalance plugin, I found that LALAL.AI stems had more definition and I could obtain better results. These does not mean the stem extraction software is perfect, there are some artifacts and the separation is not entirely accurate, therefore, among doubt, I go back to the main song track and make decisions from there. The song and the extracted stem audio files are attached, in the form of a QR code, at the beginning of each analysis.

<u>Song structure and description:</u>
The first part of the analysis is to listen to the song and identify its parts. Each part is divided in 3 levels of specification: parts, scenes and sub-scenes". The largest division is made from the parts of the song conformed by different musical ideas (melody, harmony, rhythm, instrumentation, etc.). They are given capital letters in alphabetical order, separating what we typically call verse, chorus, bridge, etc. These names have been intentionally avoided, because they can lead to confusion if the musical material doesn't fit the stereotypical characteristics these terms describe.[35]
To indicate changes inside sections, these general sections are divided in scenes. Each scene is given a number and can appear once or be repeated throughout the song (e.g. if the A section is repeated with a slight variation, the first will be named Asc1 and the variation Asc2). Finally, scenes can be divided into sub-scenes, when two scenes are very similar but only minimal details make them sonically different (e.g. a spatial or FX change) they are labeled with sub-numbers, for example, Asc 1.2.

---

[33] Audio files extracted as parts of another music audio file. Stem separation separates instruments in a track, usually it can isolate vocals, harmonic instruments and rhythmic instruments.
[34] Available at: https://www.lalal.ai/
[35] For example, the term chorus, is starting to be replaced by the term hook, since, in the last years, what we used to call "the chorus", the most climactic and explosive moment of a song, is now commonly approached contrarily, in a minimalistic, simple approach, making the word "chorus" un-appropriate. You can listen to this Switched on Pop podcast (https://switchedonpop.com/episodes/look-at-selena-gomez-now-with-justin-tranter-ian-kirkpatrick) to hear producer Justin Tranter talk about this change of approach among producers in LA.

For clarity of description, these sections have been marked inside a DAW and labeled by name and color.
All the structural information is then gathered in a table, where every element is described. The descriptions are made from only listening, and they serve to describe the sonority of the instrument, its timbre, position, and wether it has some effect, to get an overall view of all the elements that are taking place in each part of the song and where they are repeated or changed.

<u>In depth analysis:</u>
After every element in the mix has been presented, I divide the analysis in three sections: timbre, space and loudness/dynamics, as stated in Chapter 3.3. Each of the mix groups, also mentioned earlier (vocals, rhythm and music), is analyzed in each of the sections, now with the help of the graphical visualization tools, for the whole song and the stems, and the most relevant findings are described.

Lastly, I take a look to the overall sonority of the song having in mind the contribution made by the previously mentioned elements, and I elaborate a list of technical and creative post-production techniques found along the analysis.

3.7 Selection and justification of the songs under study

The songs that I picked for the analyses are all 2020's hits. The selection was made based on three criteria:

- Chart success: I wanted to work with massive hits, to unveil which features are hidden in massive songs that usually people don't experience unless applying critical listening.
- Music magazine and critic's reviews: songs that have been highly rated and which appear in top song's and albums lists on major media sources, such as NME, Under the radar, Rolling Stone, Idolator, BBC, etc.
- Mixing and mastering engineer's credits: songs mixed and mastered by globally recognized sound engineers

Unintentionally, three out of the four chosen songs are of female artists. According to Rolling Stone's ranking of the best songs and albums of the first half of 2020, female artists are greatly underrepresented. Among the Top 100, only 19 songs and 17 albums by women are part of the charts, although lots of the high-ranking positions (if we consider only pop albums) are occupied by women.

1. Physical - Dua Lipa

Produced by: Jason Evigan and Koz. Mixed by: Matty Green. Mastered by: Chris Gheringer

Released on the 30th of January as the second single in the artist's second album, Future Nostalgia, Physical is one of the biggest hits of 2020. It is a dance pop hit with 80's vibe and references, acclaimed by the critics for its constant energy, vocal performance, forcefulness and catchiness. It is a big part of the sound of 2020 around the globe, and it is an example of a clear and energetic mix.

I have chosen this track, on top of its impact in pop culture that year, because it is a great example of definition and clarity in mixing. The production is excellent, Jason Evigan uses layers of timbre to create forceful sections that sound like a whole and plays a lot with effects and instrumentation, which allows mixer Matty Green to exaggerate these traits and make every part surprising and engaging.

Mix engineer Matty Green began as an assistant for the very known engineer Spike Stent, and has worked with a variety of artists such as Weezer, Lana del Rey, Rosalia and U2 among others, proving that he is a very versatile mixer.

2. XS - Rina Sawayama

Produced by: Clarence Clarity, Shearer, Lyon. Mixed by: Tim Rowkins. Mastered by: Robin Schmidt

Rina Sawayama was an unknown London based artist until 2020, when she signed to the label Dirty Hit and released her debut album SAWAYAMA. The album was a huge success, appearing in the top album charts in renown music magazines such as NME, Under the Radar, BBC, NPR, etc. acclaimed for its innovative sound and rule-breaking fusion of pop, r&b and electronic rock, crowning Rina Sawayama as a very promising pop artist. The song "XS" was released on 2nd March 2020 as the third single on SAWAYAMA. It is a critique of capitalism in the era of climate change.

I have chosen this track because of its originality. Not only it surprised me for its mix of styles, but it is full of curious timbres, dynamic changes, atypical spacialization, etc. as well as clarity and very interesting uses of effects.

Tim Rowkins is a British mixer who has worked with artists such as Mura Masa, Arlo Parks, Two Door Cinema Club, Elton John, Rebecca Black, Bonobo, etc. he claims to be a mixer focused on detail and bringing depth and space to a mix as well as getting a big and full vocal sound. He is based in his studio in Brighton and he also does work as a producer and recording engineer.

3. Vulnerable - Selena Gomez
Produced by: The Monsters and the Stangerz, Jon Bellion, Schoudel, Gian Stone. Mixed by: Serban Ghenea. Mastered by: Chris Gheringer

Selena Gomez's latest album, Rare had been long awaited, since the artist's last album had been released 5 years before. "Rare" is a big change in Selena Gomez's style, turning to a more minimalistic and raw feeling in respect to her other work. It got great reviews on the media, (NME, Metacritic, Variety, etc.) and it topped Billboard's 200 and Rolling Stone Top 200 lists in the US. The critics claimed it was her best album yet, flattering mostly the narrative and the production.

I have chosen it mostly because of the treatment of vocals and effects. I wanted to explore how the effects were treated, because they act as main-characters of the song, which is not their usual role, and see, as well, how the vocal is treated in the different sections, and its relationship with the effects and the background vocals.

The mixer of the song Vulnerable, and most of Selena's songs, is Serban Ghenea, one of the biggest pop mixers of the last decades. He started working under producer Teddy Riley in the 90s. He had his first number one with the song "No Diggity" by Blackstreet, and is now the mixer of the biggest pop stars (Ariana Grande, Adele, Taylor Swift, Bruno Mars, etc.) currently owning 19 grammy awards and 3 latin grammys. Unfortunately, he remains a very private figure and rarely gives interviews or speaks about his work.

The mastering engineer of the song is Chris Gehringer. Ghenea and Gheringer are used to working together at Sterling Sound Studios in LA.

4. Rager teenager! - Troye Sivan

Produced by: Oscar Görres. Mixed by: Manny Marroquin. Mastered by: Chris Gehringer

Troye Sivan is an Australian singer songwriter, actor and former Youtube star. In 2013 he was signed to the label EMI Australia and in 2014 he released his first EP "TRXYE", which debuted number one on iTunes in 55 countries and reached number five on the Billboard 200 list, catapulting him to fame. Since then he has published two studio albums "Blue Neighborhood" in 2016 and "Bloom" in 2018, and several EPs This song is part of the "In a Dream" EP, which was released in August 2020.

The whole album has an electro-pop, atmospheric vibe to it and a dark, gloomy feeling. "And so exploring the sonics like that, as a storytelling tool, that's more interesting to me, rather than trying to make a cohesive pop record or something."[36] (McGregor, 2020)

I was first attracted to analyze this song because of the vocal treatment, which grabbed me from the start. In the first part of the song there's minimal instrumentation, but only the vocal effects achieve to catch your ear and keep you guessing what's coming next. When the beat and the rest of the instruments come in, the song gets bigger and we are surrounded by a space that is constantly switching between the intimate and the otherworldly. "I feel like completely besides anything going on in my life, the last few months has felt like a dream for so many people. And it's like, "Am I in a dream? Or am I in a nightmare?",' he said. "We can't leave our houses. So the thought of being able to fall asleep and go somewhere else or just any sort of escape through creativity or making something or whatever it is for you…" (The Au Review interview, 2020).

[36] McGregor, T. (20 August, 2020). *Troye Sivan's new EP In A Dream is just one of his many iso-accomplisements.* The Au Review magazine. https://www.theaureview.com/music/troye-sivan-in-a-dream/

4. Analyses of four pop songs

4.1 Analysis 1: Physical - Dua Lipa

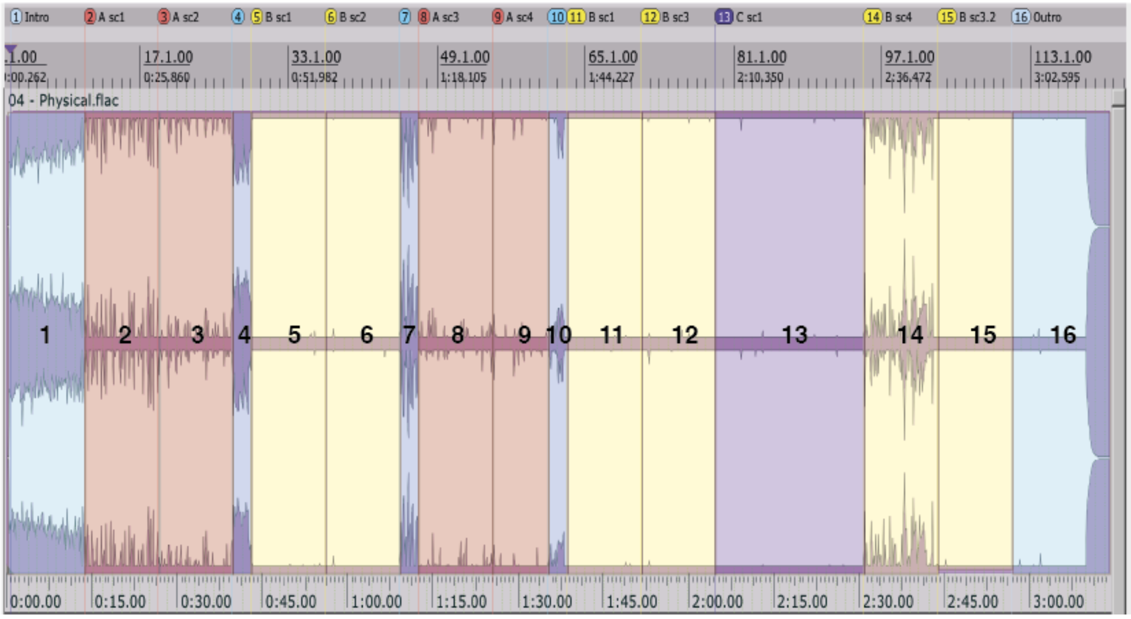

Figure 4.1.1: *Structre of Physical by Dua Lipa.*

4.1.1 Presentation and musical acknowledgements:

The song is in the key of A minor and uses only diatonic[37] harmony (Am (i) - F (VI) - C (III) - G (VII)). The time signature is 4/4 and the tempo is 147 bpm. According to Dua Lipa, it has musical and lyrical inspiration on 1980s music and films, and on an interview with Apple music, she said they created the song at producer Jason Evigan's house and that the song had an "Eurythmics vibe to it"[38], a synth pop British band from the 80s.

The song was recorded at The Bunker and RAK Studios in London as well as at Modulator Music in Toronto but the vocals were recorded at TaP Studio, also in London. It was mixed at Matty Green's Studio, Studio 55, in LA and mastered by Chris Geringer at Sterling Sound, New Jersey.

[37] Harmony derived from the key of the song.

[38] Apple Music Album Release (22 April 2020). *Future Nostalgia by Dua Lipa*. Retrieved 22 February 2022, archived from the original 22 April 2020.

As most pop songs, the leading element of Physical are the vocals. Dua Lipa is a mezzosoprano[39] singer. In this song, she exploits most of her vocal range (the lowest note (excluding the E3 on 1:14 which is almost vocal fry) is A3 and the highest is C5) and uses a variety of vocal techniques, for example, adding air to the end of sentences, almost vocal-frying[40] on the breaks, subtly tearing her voice on some words (e.g. 1:43'), etc. making it a very dynamic vocal performance. Her voice is slightly boxy (around 500-700Hz), a feature that hasn't been taken away in the mix because it's part of her vocal identity. The low end of her voice is probably her strongest trait and is key to delivering the power on her performance, although in the mix, the high end is more enhanced.

The second most important element in the song is the beat. From the beginning it is presented very tight and "in your face", dry and loud, and it stays this way throughout the whole song. It is clearly not played by a human and despite the fact that the sound is modern instead of 80s (there are no big reverbs and big overhead splashes) the pattern in conjunction with the bass brings out this updated 80s feel[41].

Looking at these two main elements on Physical, beat and vocals, we can see that the rhythm is very straight, always falling on the downbeat, whereas, in opposition, the accents on the words are constantly on the upbeat. This gives the song a lot of movement and everything else works to emphasize this sensation and achieve what is the primary intention of the track, making people move.

---

[39] The register of a mezzosoprano singer goes from A3 to F5.

[40] Vocal technique consisting of obstructing the airflow on the glottis, producing the lowest human vocal register.

[41] For comparison, listen to the chorus on "Here comes the rain again" by Eurythmics (which is curiously in the same key) for almost the same bass-beat relationship on the chorus.

4.1.2 Structure:

1	0:16	Intro	There's a sequenced synth pattern that fills the mid frequencies and stereo flutes with reverb on the high end. A band-pass filter is applied to all the instruments from approximately 200Hz to 2k with a broad slope (probably 6dB/octave). The filter gets opened gradually as the part ends.
2	0:29	A sc1	The mix gets organized in 3 clear layers, from front to back: beat, vocals and bass. The beat is upfront with no reverb. Kick drum lowest resonance is at 60Hz and snare drum fundamental is at 130 Hz. Every two measures, the snare drum is mixed on the first measure with a clap sound and on the second measure with a high snare sound bathed in reverb. There are overhead sounds located widely at left and right, playing with panning and various sized reverbs, making them appear in a different space in relation to the beat. Vocals have slap delay and medium reverb with pre-delay, to make it sit upfront. They are placed behind the beat, but they are present and full. There are also vocal chops in the background doing melodic details. Bass does a volume crescendo to fade in the next part. There is no sub information yet. Instruments: high pitched vocal chop and flute-like sounds, both bathed in a big reverb, that puts them behind on the mix. A sequenced synth pattern fills the midrange and some of the low-mids of the spectrum.
3	0:42	A sc2	Beat stays the same Bass sub lows increase in volume, there's no bass information above 2k. It is not completely mono, it is constructed with layers and some of them fill some of the stereo space, because they have a short reverb and some modulation effect, probably chorus.

			Main vocal FXs stay the same and vocal harmonies come in very wide on the stereo field and thinner than the lead (with less low-end).
4	0:45	Break sc1	Vocals become completely dry and a clock ticking sound functions as a four on the floor kick pattern.
5	0:58	B sc1	Beat: Changes completely. Kick loses attack impact. Snare changes, the fundamental is still at 130Hz but more layers are added and overall it has a higher pitched sound, with resonances at 1300Hz and has obvious reverb (not big but present, probably room reverb). There's a new clap sound and pattern and the hi-hats enter here, doing a swinged semiquaver pattern, accentuating the upbeat to achieve a funky type of vibe, and it is in the middle. The sound occupies the highest end of the spectrum. The general spectrum of the beat has an obvious dip in the mids (200Hz-2kHz). Bass: The spectral information in the low end stays the same, but it has less attack, probably due to the reverb that comes in at this point, that softens it, and lots of sustain. It is also made wider, and the modulation effect is made more obvious. At the end of the passage, it does a crescendo. Instruments: Synths, bass-like sequence, strings and synth vocals. They spread over most of the spectrum (all the low-mids to high-mids) and add the sense of space (big reverbs and wide stereo image) and movement (the vocal synths are automated in panning). In this part, synth vocals are added (vocoder-like sidechained background vocals) and along with the synth keys, they form a texture on their own, and are a bed of sound for the strings, which sit on top. The bass sequence is behind them, lower and in its own space. All elements together have, at the same time, a quarter note and a half note sidechain feel. Vocals: The vocals are double tracked, and have 3 delays; a short slapback delay, to lift the vocal, a quarter note delay with reverb and sent to the back, which adds space and widens it, and a whole

			note delay, filtered (mid-high sounding) and automated, appearing only in some words, being used as a response to the main melody. All the delays together give the impression of spatial depth to the vocal. For background vocals there's a male vocalist doing the melody an octave down and crowd-shouting vocals at the end of the passage.
6	1:11	B sc2	Everything stays the same except new background layers of synths and vocals appear and the overall texture gets fuller and more open.
7	1:14	Break sc2	The kick drum sound changes for a more boomy (60Hz), less attack sound and no overheads. We can hear the tail of the reverb and delays of the vocals' fades during this part and a sine wave type of sound doing a glissando during the first bar that fades in volume.
8	1:27	A sc3	Very similar to A sc2 but the stereo image is wider because of the bass and sequence pattern become more stretched and chorus-like.
9	1:37	A sc4	Hi-hats at sixteenth note pattern enter here (the same we had on the B sc1). A new synth sound appears, filling some more of the 2.5kHz to 6 kHz area. It is placed slightly to the right, and it is placed back in the mix.
10	1:40	Break sc3	Same idea as in Break sc1 but now some of the instrumentation is kept. The bass, the sequence and the hi hat keeping the beat. On the 2nd bar the bass is muted, and the sequence is filtered and distorted with a riser underneath to go to the chorus again.
11	1:54	B sc1	Same as B sc1.
12	2:07	B sc3	A new synth, with a triangular/saw wave type of sound, comes in doing chords and adds high end. New layers of background vocals are also added. The bass continues the same but has a "cassette tape stop" effect at the end.

			At the end of the part, the reverb on the drums, especially the snare drum, is automated to be bigger and obvious.
13	2:33	C sc1	The instrumentation on part C is the same but vocal melody changes. Bass comes in with the same cassette effect but backwards ("restart" effect). The synth added in the latter part is made even more important because it does higher pitched chords. Vocals here reach the highest register. Dua Lipa's voice is not as powerful in this register, so the vocal melody is doubled by L-R background vocals that back it up.
14	2:46	B sc4	This part works as an extension of the bridge (C) but musically it has the same ideas as the chorus. It is narrowed to the vocals (with the background louder and very wide) and the dry beat (kick, snare, and claps).
15	2:59	Bsc 3.2	Everything comes in again for the final chorus, and some vocal harmonies are made louder than B sc3.
16	3:13	Outro	The outro is an instrumental of the chorus that ends abruptly leaving only the tails of reverbs and delays of all the components.

4.1.3 Frequency/Timbre

This song serves as a great example of timbre layering and frequency separation. The selection of the sounds is key to this track, each instrument is carefully chosen and fused, to create a more complex texture. We can categorize the instrumentation in four big groups: bass, beat, music and vocals. We can analyze each of these groups to get a full understanding of how the spectral energy is distributed in order to avoid frequency masking between them and achieve clarity.

The low end of the beat is defined by the kick and the fundamental of the snare, and the high-mids and highs are where the overheads live. The beat has a clear dip in the mids (300Hz-1kHz) (Fig. 4.1.2) that gets even more dramatic during the choruses (Fig. 4.1.3), probably done to make space for the other instruments.

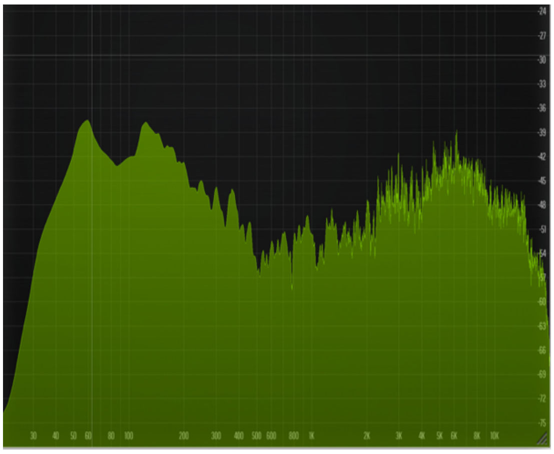

Figure 4.1.2: *Spectrum of the beat on Asc1.*

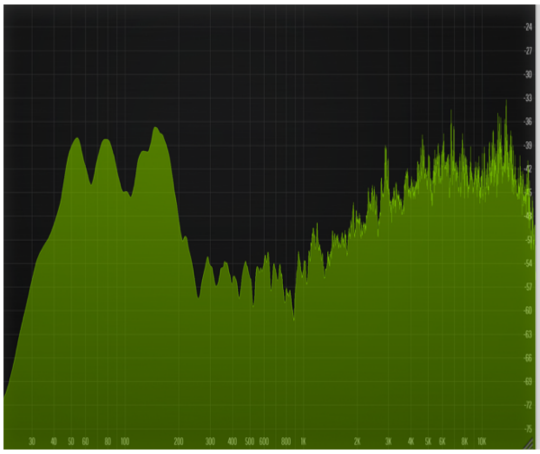

Figure 4.1.3: *Spectrum of the beat on Bsc1.*

In the chorus and the bridge there's a lot more high-end presence from the beat because a brighter hi hat is added, which makes the chorus sound more open and shiny than the verse. Contrarily, the kick drum main resonance moves from 60Hz on the verse, to 50Hz (with also a lot of presence of 75Hz) on the chorus, adding energy to the low end, making the overall sound on the chorus more extreme, exaggerating the limits of the frequency spectrum.

The bass sound is very round, it has little attack and lots of sustain, creating a sound as if the notes merged with one another and it also doesn't affect the attack of the kick drum. The fundamental of the bass is around 50Hz (depending on the chord, the lowest fundamental in the song is 45Hz, with lots of energy even at 30Hz), carrying the sub-low energy all along. In the chorus it is a bit brighter due to the modulation effect and the reverb which are added. The bass and percussion together continue having a dip from 200Hz to 1k. This dip is filled by the harmonic and melodic instrumentation.

We can see on Figures 4.1.4, 4.1.5 and 4.1.6, during the verse, the music doesn't have a lot of elements. The arpeggiated verse is the main element and other punctual details go in and out, but the main energy is between 200Hz and 2kHz. Then, on the chorus, the spectrum wideness and the high-mids are filled and spread along the stereo field (Fig 4.1.7). The brightest part of the music is the bridge due to the very bright synths and strings.

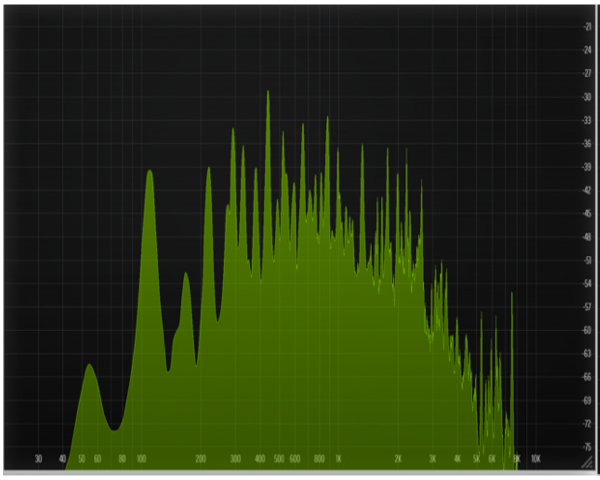

Figure 4.1.4: *Music group spectrum on Asc4.*

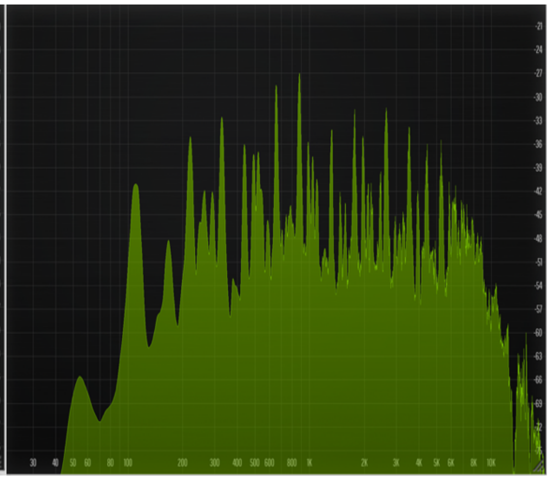

Figure 4.1.5*: Music group spectrum on Bsc2.*

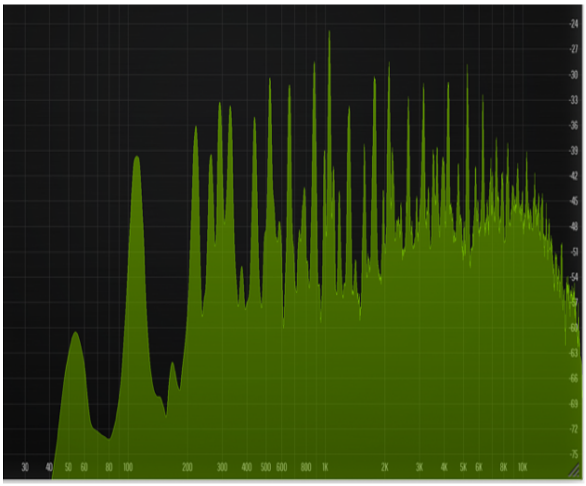

Figure 4.1.6: *Music group spectrum on Csc1.*

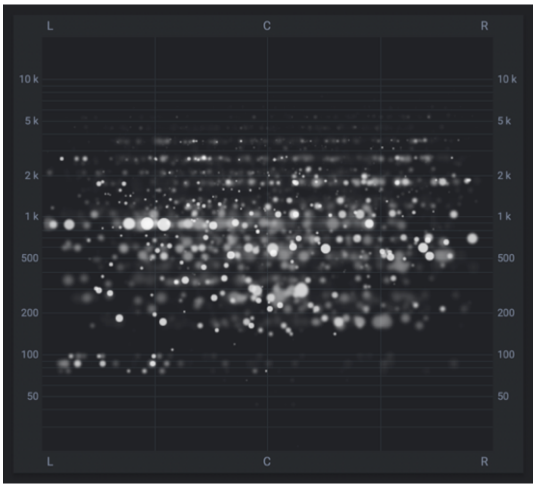

Fig 4.1.7: *Stereo spread of the spectrum of the music stem.* X axis=stereo field, Y axis= frequency bands plugin by

If we focus on the vocals, we can differentiate timbres between the verse and the chorus. In the verse, (we can hear the verse's vocal timbre clearly and dry on the breaks), there's emphasis around 400Hz (because it is the fundamental frequency), which makes her sound close and warm, and we can perceive brightness and air. In the chorus, the warmness is removed and high-mids are maximized, the vocal sounds more defined and shinier. In this song, the most crowded section is the last chorus, as we can see in Fig. 4.1.8, the whole audible spectrum is being used. In this situation, it is important, while mixing, to define which elements are more important and should remain over others. In this case, the beat and the vocals are the most consistent element of the mix and everything else is carefully sculpted to fit the space around them. The harmonic and melodic details are most likely to stand in the same frequency bands as the vocals so, in this song, the music is very widened, so that its high-end energy will not mask the vocals brightness, which is kept in the middle.

The overall shape of the spectrum is flat, with a dip between 200-400, probably to avoid muddiness.

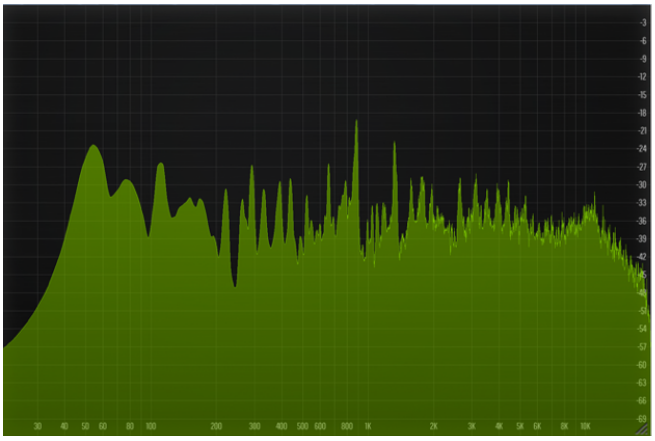

Figure 4.1.8: *Average Spectrum of last chorus Bsc3.2.*

4.1.4 Space

Like in a lot of popular music songs, the main elements are placed in the middle for the whole song, but this mix doesn't feel mono at all, moreover, it uses space in a dynamic way.

Using a Lissajous meter (or also called, Goniometer) we can observe (Fig 4.1.9) that the mix widens as time goes by. The second verse is wider than the first, the chorus incorporates countermelodies and background vocals that widen the image, and the bridge is the climatic part, where the background vocals and strings open the image even more.

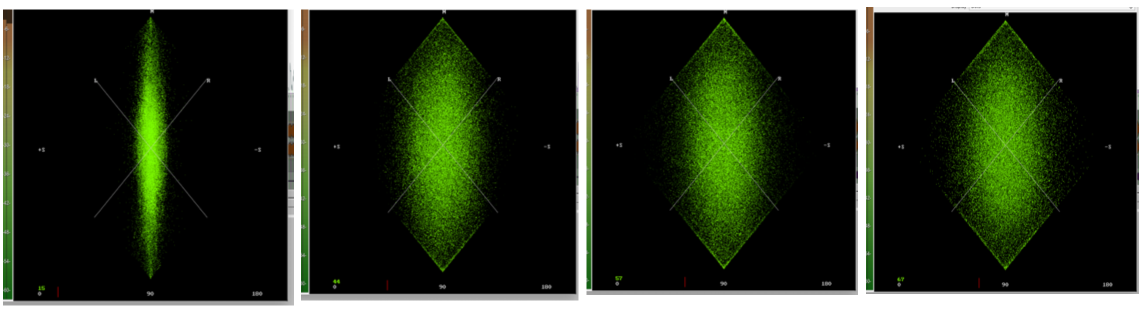

Figure 4.1.9: *Lissajous meter display of stereo field during verse, chorus, bridge and outro.* Source: Screenshot of Reaper Goniometer.

The bridge is also a great example of the layers of the song. If we listen to it in mono, we can clearly identify what is in the forefront: the vocals and the drums. We can hear the arpeggiated bass pattern on a secondary plane and strings and background vocals further in the back. Listening in this way makes us understand how many elements have a play outside the middle, and how the song doesn't work the same when we take them out, even though the main elements are there. When we hear it all together, we can perceive how the vocals and beat are as present and defined as in mono, but now surrounded by all these layers at different depths that don't clash with them but make them richer. The mixer (guided by the production decisions) is creating movement and unconscious flow for the listener by playing with the stereo image and bringing changes to every part, and, as we have seen, using the stereo spectrum to the fullest helps with frequency separation because close frequencies don't fight for the same space, a resource used mainly with the harmonic content and the vocals; by bringing the music to the sides, the vocals sit in the middle and the timbres don't get compromised as much.

In fig 4.1.10, it is displayed how both, the stereo field and the frequency spectrum are used to the fullest in this song, but thanks to the layering it is not perceived as a wall of sound coming from all sides. Every sound has its place and contributes to the general soundfield.

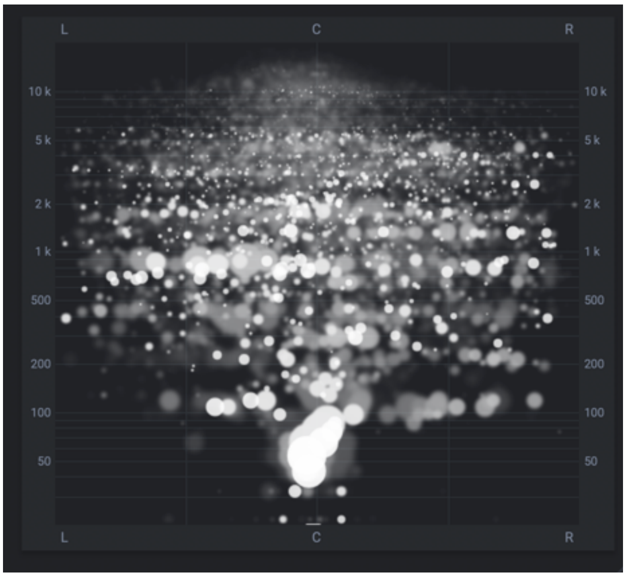

Figure 4.1.10*: Stereo field and frecuency spread on last chorus (Bsc3.2).*
Source: Screenshot of Zplane Peel plugin.

4.1.5 Loudness/Dynamics

The Integrated loudness of the song is -6.8LUFS, which is loud for any streaming service and platform, and the Loudness Range sits between 4,5 and 5 LUFS, which means that the track is heavily limited and compressed. For pop mixes, a higher than 5LU loudness range is advisable to not make the song completely static but maintain consistent energy[42]. If we import the percussion track into Izotope RX's waveform stats, we can see the Loudness Range of the beat is nearly 2 LUFS. It is the element with the lowest dynamic range, which means it has been very compressed intentionally, to bring consistent energy to the track.

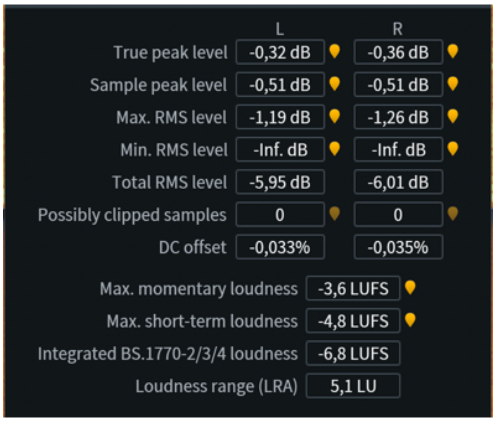

Figure 4.1.11: *Loudness measurements on Izotope RX 8 Waveform Statistics*

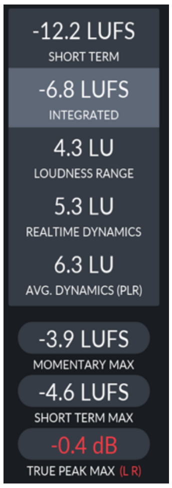

Figure 4.1.12: *Loudness measurements on Youlean Loudness Meter.*

On the whole, we can see that dynamics are brought to the song thanks to the arranging and structure, more than the performance and treatment of each instrument. The song is very energetic but it doesn't get tiring because it has breaks where the ears can rest and the duration of the sections is short enough to change into something sonically different every time we could start to get tired and, in that way, refresh our ears.

For a very compressed mix I believe the engineers achieved to maintain a sense of depth and space that make this song interesting and keep us engaged until the end.

---

42 Introductory guidelines on loudness metering and mastering can be found in: https://www.masteringthemix.com/pages/mixing-with-levels?currency=EUR

The whole Future Nostalgia album has very little dynamic range, as we can see in fig. 4.1.13. This happens with most mainstream pop albums today, mastered with radio and clubs in mind, instead of an experience of the album as a whole.

DR scale

00 01 02 03 04 05 06 07 08 09 10 11 12 13 14+

Artist	Album	Year	avg DR	min DR	max DR	Codec	Source
Dua Lipa	Future Nostalgia (The Moonlight Edition)	2021	05	04	06	lossless	Download
Dua Lipa	Dua Lipa	2017	11	10	11	lossless	Vinyl
Dua Lipa	Future Nostalgia The Moonlight Edition	2021	10	09	13	lossless	Vinyl
Dua Lipa	Future Nostalgia	2020	05	04	06	lossless	Download
Dua Lipa	Future Nostalgia (Yellow Cassette)	2020	11	10	13	lossless	Unknown
Dua Lipa	Future Nostalgia	2020	07	07	09	lossless	Vinyl
Dua Lipa	Future Nostalgia	2020	05	04	06	lossless	CD
Dua Lipa	Future Nostalgia	2020	05	04	06	lossy	Download

Figure 4.1.13: *Dynamic Range chart of Dua Lipa's albums.*

Source: Dynamic Range DB. https://dr.loudness-war.info/album/list

4.1.6 Conclusive thoughts

In pop music today, the final result comes highly defined by the production. In this case, the mix is highlighting how the production suggests movement and space, creating a dynamic mix that uses the frequency and stereo spectrum to the fullest, and achieves a powerful mix while maintain layers of depth.

Some of the ideas I have taken away from this song are:

- Using the song's structure to create dynamic changes.
  - Changes of stereo wideness and element placement.
  - Changes of Fx on differrent parts (such as vocal reverbs and delays on chorus vs bridge).
  - Enhance changes in timbre of the same instrument in different parts (e.g. sound of kick drum or bass on verse vs chorus).
- Using the stereo field to spectrally separate similar instruments/parts.
- Using chorus on low-end instruments (e.g. bass) to make the mix wider and leave space for the kick drum.
- Interesting FX details like:
  - Automating an increase of the reverb of an instrument on the last bar of a part to transition into the next.

- Using more than one sidechaining pattern in different instruments to achieve movement and groove.
- Taking completely away the reverb (or other FXs) on a phrase or break to get attention.
- Filtering the high end on background vocals to make the lead vocals more present.

4.2 Analysis 2: XS - Rina Sawayama

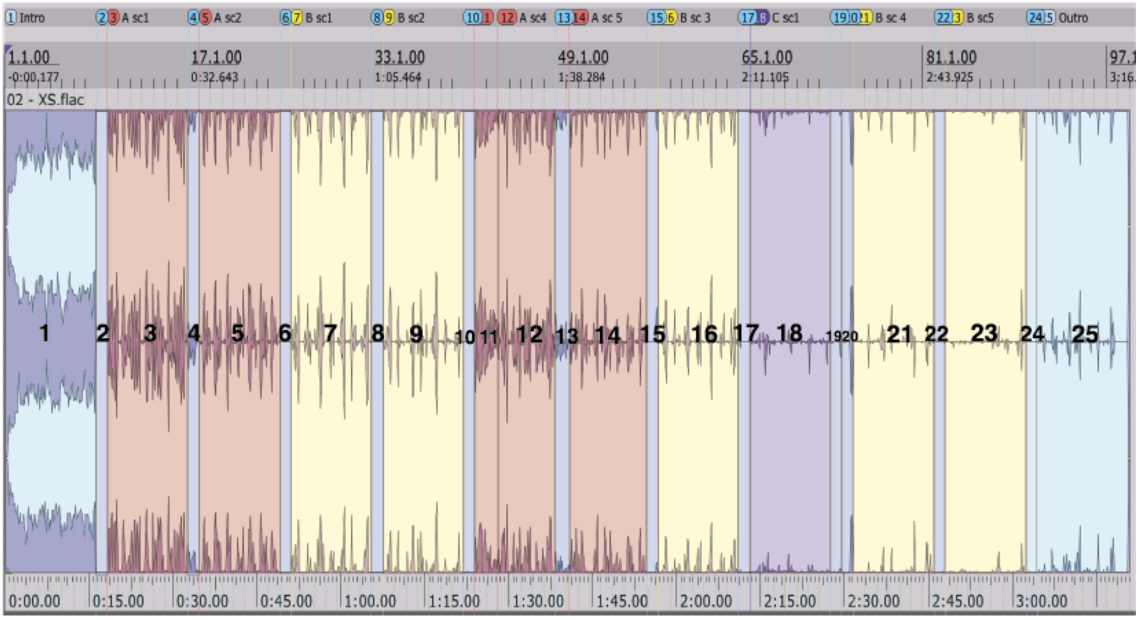

Figure 4.2.1*: Structure of XS by Rina Sawayama.*

4.2.1 Presentation and musical acknowledgements:

From the first time we listen to this song, we can see the interest here relies heavily on the production. Three producers worked on this song, Clarence Clarity, Shearer and Lyon. Clarity is a singer songwriter who had worked on Sawayama's first EP and contributed with songwriting and producing in most of the songs on this album. Kyle Shearer is a well-known songwriter and producer who has worked with big pop acts such as Carly Rae Jepsen, Tove Lo, Melanie Martinez, etc. and only has credits for this song on the album. The third producer is Steve Lyon, who is an experienced music engineer specialized in alternative and electronic rock. He has big credits like Depeche Mode, The Cure, Recoil, etc. and has also worked as a producer for many independent rock artists such as How To Live and Suzerain. From these 3 different types of artistic contributors, we can understand dynamism and fusion were created.

"XS" is in the key of D minor but the whole song is built upon 3 chords: iv (Gm), V7 (A7) and bVI (Bb). It is a 4/4 beat at 117 bpm, with lots of syncopation on the percussion

and the vocal layers, making it a really fun song filled with ear-catching details and sudden changes.
Mixer Tim Rowkins is used to working with indie pop artists, he has worked with Mura Masa, Arlo Parks, Two Door Cinema Club, etc. and had mixed Rina's first EP "RINA". The mastering engineer is Robin Schmidt, a Grammy-nominated engineer who has worked with many mainstream pop and rock artists from his studio "24-96 mastering" based in Germany.

4.2.2. Structure:

1	0:16	Intro	Tape emulation type of effect. Ambient car sounds. Strident strings and vocal chops (the sound reminds of a Chinese opera). The stereo correlation is close to 0. It is the quietest part of the song, at -12LUFS.
2	0:18	Break sc1	Very loud change, from -12LUFS to -4.5LUFS (short-term loudness measure), and true peak max 0.3. Very wide distorted electric guitars (negative correlation). Kick drum, and very low toms hit in the middle while overheads splash the stereo. The drums make a typical hard-rock fill which sounds unexpected and out of context in the song. The overall sound is very compressed.
3	0:32	A sc1	Percussion: Kick drum fundamental 60Hz, claps from >1kHz, they are bright, panned to L-R and have a very short reverb. Toms sit above the kick drum spectrally (±150Hz-500Hz) and they are tuned, making a descending melody. They have a different sound to Break sc1 and each one is panned differently and they are wider and closer than the claps. They have a bit of reverb which contrasts with the dryness of the kick drum. There's a shaker doing a syncopated pattern spectrally above the claps and a chime sound that's very bright and dry in the middle.

			At the start of the 4th bar, the shaker and clap are suddenly cut for a very brief moment, then the toms hit again. On the 5th bar a bubble-like sound motif comes in and out in the middle of the stereo. The harmony is carried only by the acoustic guitar. It sits in the middle and has a small reverb that widens it a bit to the sides. The guitar sound is in the high-mids area and played with a pick. Vocal sound is pretty dry, and emphasized at 1-2kHz which makes it sound a bit narrow, and it also has a lot of air (>10kHz). It has a small bright reverb with early reflections that give a very subtle slapback.
4	0:34	Break sc2	Timpanis are tuned to G and D. Kick drum does a crescendo. Vocals have a different timbre, 2k-6k are filled and 1kHz is dipped. The reverb gets louder and makes the vocal brighter and more open. Bright Harp-like sound appears (maybe a Koto), mostly on the sides, for the change of section.
5	0:49	A sc2	Perc: same without bubble and silence details and the shaker starts a more active pattern. The chime percussion is substituted by a chime synth, which is bright (>1kHz) and wide, sent back with a big reverb. The bass enters in this section. The fundamental is around 50Hz, it's a very subby sound, doesn't have attack. The bass is not completely mono and the pattern follows the kick drum, and is under it in frequency. Guitar: Same as Asc1. The vocals have more body (more energy around 200Hz) and brightness, more high-mids, probably helped by the change of reverb, which is mid-long with lots of early reflections, high passed and louder, adding an obvious and bright slapback in stereo, widening the lead a lot and making

			it sound "spacey". The reverb time is intended to decays just before a new sentence comes in. Sibilance has been removed around 6-7kHz.
6	0:51	Break sc3	Similar to Break sc1. The drums are a bit softer in respect to other breaks, maybe because there are vocals and they need to be understood. The vocal timbre is metallic, it has a modulation effect similar to the comb filter sound.
7	1:05	B sc1	Percussion: Same kick and snare drum sound but now timpani follow the kick drum pattern, adding a layer of texture. There are more overheads, hi hats syncopating details and the shaker changes pattern to a more continuous one and it is brighter. Bass: Probably a new layer of bass is added because it has more attack and is louder. Guitar: same as above. Lead vocal: it is doubled and has the same timbrical qualities as Asc2. Background vocals enter in this part. There are 3 types ("more", "excess", "oh me, oh my"): "more": sounds close, wide but slightly panned to the left, dry and with more body and less brightness than the lead (it is probably low-pass filtered to differentiate it from the main vocal). It sits behind the lead in a secondary plane. "excess" very bright (has a lot of energy at 10kHz) high-pass filtered and sibilant, it sounds like a whisper. It is the one that sounds the furthest. It has a variation with a panning automation to hard left and right at quarter note tempo on the 5th bar. "oh me, oh my", it is the only background that is mono. It is pitch shifted down and in the same plane as the lead.

			The background vocals have a lot of contrast with the lead, timbre, rhythm and spatial characteristics are contrary to it creating a kind of call and response feeling.
8	1:07	Break sc4	Same as Break sc1. Loudest break so far. It doesn't have vocals.
9	1:21	B sc2	Drums: More hi hats are added, with a trap-like sound and an open hi hat on the up-beat every 2bars. More claps are added too. Guitar and bass are the same as B sc1. Lead vocal is the same as Bsc1 and a new background is added; a vocoder-like vocal, harmonizing the lead. It is bright and wide.
10	1:23	Break sc5	Same as Break sc1.
11	1:28	A sc3	Simplified beat, only clap sound + coin fx. Bass doesn't play. Guitar: same. Vocals like sc1 with more obvious reverb. A new background vocal harmonizes it. It is darker than the lead but sits in the middle behind it, sent to the back with a bigger reverb than the lead.
12	1:38	A sc4	Percussion: like Asc2. Perc chimes come in again. Bass and guitar: same as Asc2. A string synth sound comes in (similar to a clean electric guitar note). It is a pedal note that moves automated in a quaver-triplet pattern from left to right. Its timbre is narrow, between 2 and 4kHz. Vocals: Same as Asc3. Background vocals come closer than Asc3 and serve as a harmonic pad.
13	1:40	Break sc6	Same as Break sc2.

14	1:54	A sc5	Percussion: It doesn't change except some shaker variation/details. Guitar and bass are the same. Strings: They occupy the high-mids are and are very stretched in the stereo. They occupy the position of the pedal synth at Asc4. They are in 2nd plane. Background vocals: new harmony vocal, in the middle, behind the lead but darker. Same situation as Asc3. Detail backgrounds saying "aha" are added and occupy 1kHz-2kHz on the sides.
15	1:56	Break sc7	Same as Break sc1 + same vocal effect as Asc3 + harmony vocal panned right.
16	2:11	B sc3	Perc: like Bsc2 + more variations on clap and hi-hat. Strings continue like Asc5. From the 5^{th} bar higher octave strings come in. A new background is added, doubling the lead an octave higher.
17	2:13	Break sc8	Same as Break sc1.
18	2:27	C sc1	Bass: more sustained sound with more sub-bass energy and a bit distorted. Percussion: The sound is the same but there's a new pattern. The clap is sent further with reverb and the hi hat is less bright. There are a lot of overhead and down-riser sounds filled with reverb that create ambiance. Lead vocal doesn't change but backgrounds have more importance and reverb. The pitch shifted harmony is filtered (100-1kHz) and has more reverb. Strings New Synths: • Choir-like sustained pad on stereo and sent to back with reverb. • Bell sound on the middle. • Arpeggiated pizzicato strings on sides, dry.

19	2:29	Break sc9	Same as Break sc 1.
20	2:31	Break sc10	Same instrumentation as last break but there's a melodic effect of pitch decay. On the last sentence the vocal is left dry with a dry breath underneath.
21	2:45	B sc4	New synths: • Asc2 chimes but brighter and more present. • Harp sound comes in again panned right. • Strings make shiny pad behind in stereo. • Electric guitar line (2:40'). • 2:45 descending harp panned right. Vocals: same as Bsc3 + vocal harmony and lead vocal melody details. Hi hats 16th note continuous pattern until the end.
22	2:48	Break sc11	Same as Break sc1.
23	3:02	B sc5	New synth: • Siren-like synth doing melody line in the background, bathed in reverb. • More harp activity. 2:56' accent on clap. Further background vocal additions, panned every 1 bar L-R doing counter melodies.
24	3:04	Break sc12	Same as Break sc1 + harp descending chromatic motif.
25	3:21	Outro	Continues like Bsc5 + high-pitched string long notes. Ends abruptly with break sc1 and sudden dry mouth sound.

4.2.3 Frequency/Timbre

In this song the percussion is the element that catches our ear from the start. There are a lot of things going on in the percussion pattern and the spectrum gets filled only by this element. The lowest part is covered by the kick drum, low mids by the timpani, mids by the snare, high-mids by the shaker and the high end by hats, chimes and other cymbal elements. The pattern is very active and contrapuntal, and the kick and timpani almost create a melodic line, which, alltogether, creates a very dynamic beat and a constant moving frequency spectrum. The beat is not the same throughout the song, as it goes on, details are added, filling the spectrum even more on every part and making the song more intricate. The several high frequency elements that play altogether, more towards the last choruses, are separated in space and frequency. The shaker is the one that takes up more space and is more present, and the hi hat details are not placed in the middle and contain mostly high-end information, to make them shine when they appear and then disappear without playing over the shaker pattern.

The bass is round and fills the low end, except for the choruses where it has more attack, even though it always stays under the kick drum in frequency. We can see, on the choruses (Fig 4.2.2), that the bass has a dip around 70Hz and the kick drum (Fig 4.2.3) around 100Hz which means that they are probably sidechained to one another, in order to make space and make them work together. The bass is not completely mono, which helps with separation from the kick drum, and overall, the bass sounds softer and the kick is punchier and louder, making it stand out more.

Figure 4.2.2: *Bass frequency spectrum on the chorus*. We can see a bass dip around 70Hz.

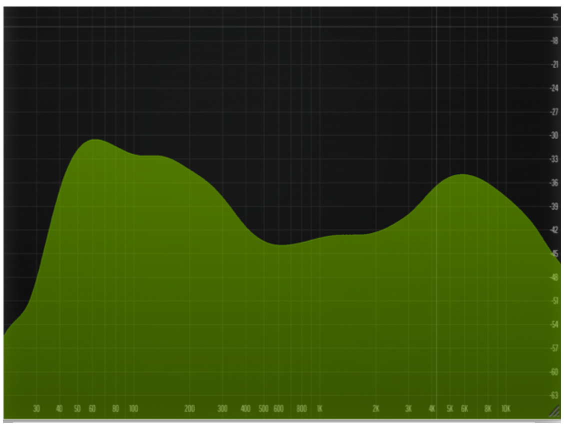

Figure 4.2.3: *Drums frequency spectrum on the chorus*. We can see dip around 100Hz.

On top of the rhythmic section we have the harmony and melody, which at first is only supported by the guitar and harp, and afterwards, more elements keep being added as the song progresses. The bridge and the last choruses are the most crowded parts. As we can see in Fig 4.2.4, the instrumentation is very wide in this part and fills, mostly, the low and high mids area. It is interesting how the spectrum information of the music stays below 10kHz, to not take away the air and shine from the lead vocal.

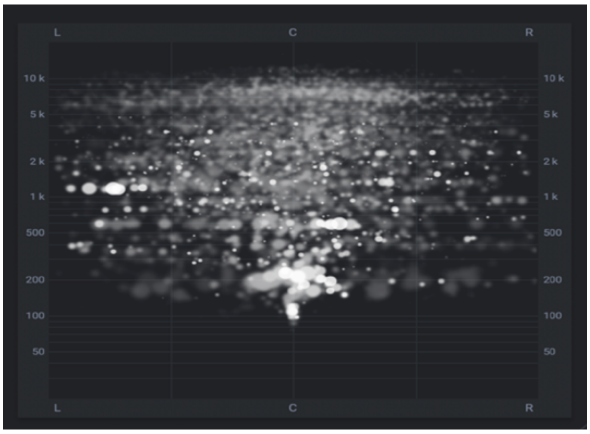

Figure 4.2.4: *Frequency and stereo placement of music stem on Bsc5.*

The vocals have different timbrical scenes. On the verses there is a curious pattern that is repeated before and after the first chorus, regarding the vocal timbre. On Asc1 and Asc3-4, we can hear the vocal has a 1-2kHz emphasis that makes it sound a little bit nasal and closed (Fig 4.2.5), but then on the following part, Asc2 and Asc5 respectively, the vocal gets opened spectrally, and apart from the reverb effects that make it lusher, we can hear more high-end energy, around 4 and 5kHz, which makes the vocal more detailed, bright, open, breathy, etc. (Fig 4.2.6) whilst still maintaining the body it had but without sounding nasal. The vocal on Asc1 and Asc4 is also a little bit distorted, probably using a harmonic exciter, which contrasts with the other verses where the vocal is clean.

Figure 4.2.5: *Vocal spectrum Asc1.*

Figure 4.2.6: *Vocal spectrum Asc2.*

On every section we can see a dip around 6kHz which is probably sibilance elimination (the song is full of esses).

The reverbs on the vocals are generally bright and have a lot of early reflections, which add high-mids energy to the spectrum. If we listen to the vocal track soloing the sides, we can hear how the lead vocal reverb changes each section. On the first verse it is shorter and more neutral, not so obvious, whilst on the second verse it gets bigger, louder and more colored (more reflections) with a very audible slapback delay. The bridge and choruses have the same reverb, a big-sized room with bright reflections. Probably all the reverbs for the lead vocal have been high passed, to avoid muddiness.

The average of the spectrum on the last chorus is shown in Fig 4.2.7. It has emphasis of the high frequencies and decays slightly and evenly towards the lows.

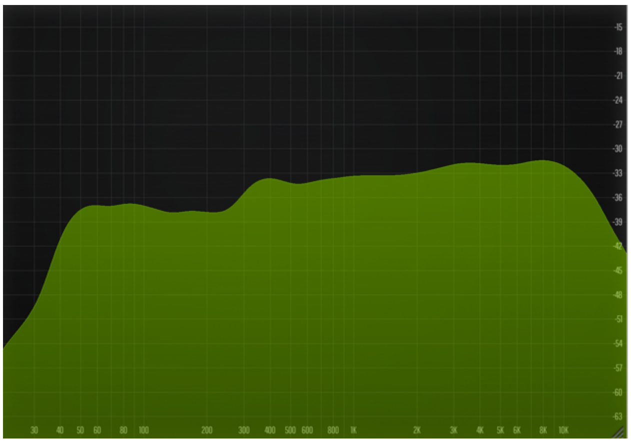

Figure 4.2.7: *Last chorus Bsc5 spectrum.*

4.2.4 Space

This is a song of extremes and contrast. The song starts with a stretched stereo image, the correlation meter flows around +0.5 sometimes going into the negative side, and then the first break comes in stretching the stereo even more and suddenly distorting everything. It contrasts with the verse, which is fairly mono (the toms and bass help to make it a little bit wider). Then, on the choruses, the background vocals are very wide and function as counter melodies from the lead vocal, which creates a back and forth flow between mono and stereo that keeps the song interesting. The mix is built to make the song unpredictable; some phrases are wide, some lean sideways, some are pan-automated… The stereo field is used to keep the ear engaged. As the song progresses, more elements are added to the stereo; background vocals, strings, synths, etc. On the chorus the bass is made wider, with a subtle chorus-like effect, which creates a more ambient sound. All this instrumentation is added in layers, but it is probably due to heavy compression that these layers are hard to perceive with definition. We can hear two clearly separate layers throughout the song; the vocals, beat and bass form the closest one and the rest of the instrumentation form the second. On the second layer, everything feels crowded and close together, like it has been separated in the stereo field but not in the depth plane, so details are a little difficult to assimilate.

4.2.5 Loudness

The integrated loudness is -6,5 LUFS with the True Peak level at +1dB (which means the song is clipping in some points) and the maximum momentary loudness is -2,7 LUFS on the 3rd break. As we can see, this song is really loud. The Loudness Range is 5 LU (measured with Izotope RX8) and 4,4 LU (measured with Youlean Loudness Meter), which means dynamics are very reduced, which explains why our ears are tired by the end of the song. We can clearly see this lack of dynamics in the shape of the waveform (Fig 4.2.8). During the first half of the song (until Asc3, leaving the intro out of the measurement) the LRA measure is 5 LU, whereas in the second half, it is 3,4 LU. The bridge, which is usually a softer part in pop songs, a release before the last big chorus hits, in this case is not softer than the choruses before it: they both measure around 6,5 LUFS of integrated loudness.

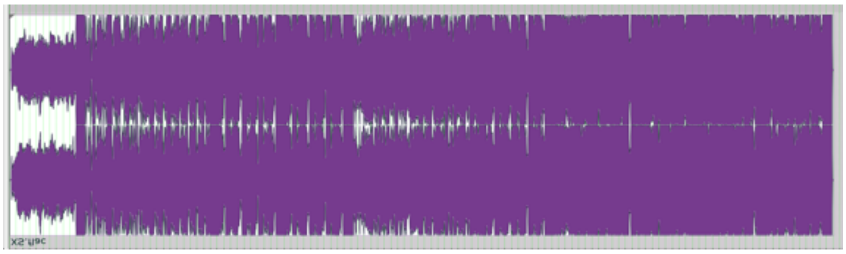

Figure 4.2.8: "XS" Waveform.

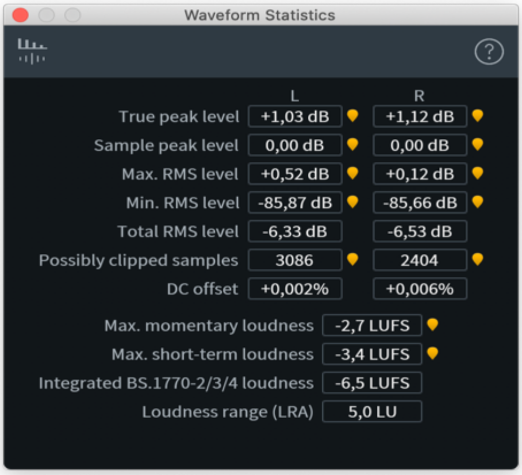

Figure 4.2.9: *Loudness measurements of "XS" on Izotope RX 8 Waveform Statistics.*

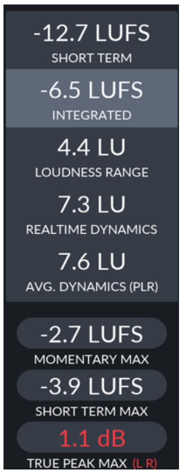

Figure 4.2.10: *Loudness measurements of "XS" on Youlean Loudness Meter.*

In conclusion, this song is very compressed and even clipped, which I believe is probably done with intention. It was probably the artistic view of Sawayama and her team to crush every sonic element of the song, make it unpredictable, and go beyond the boundaries of loudness as an artistic statement, to make everyone listen and pay attention to this artist's first album. I first thought this could be the reason, but I dug further, and I was surprised to find that the average dynamic range of the whole record (Fig. 4.2.11) is 4dB. This means that there's only 4dBs of release through the whole album.

In fact, "XS" has the 4th lowest RMS level value from the album. As a result, it would probably exhaust our ears to listen to this album on one sitting (more so if done with headphones).

Artist	Album	Year	avg DR	min DR	max DR	Codec	Source
Rina Sawayama	Sawayama	2020	10	09	12	lossless	Vinyl
Rina Sawayama	Sawayama	2020	04	03	06	lossless	Download

Figure 4.2.11: *Dynamic Range chart of Rina Sawayama's albums.*

Source: Dynamic Range DB. https://dr.loudness-war.info/album/list

4.2.6. Conclusive thoughts

"XS" is a song of contrasts and statements. The production is very playful and is constantly looking to surprise us. The mix magnifies the details and the aggressiveness making it sound "in your face" and creating a piece of music we can't predict.

Some techniques learnt:

- Eq (or dynamic-eq/sidechain) the bass and kick drum in complementary frequencies, to avoid masking.
- Using the bass to highlight the structure. In the verse it is less present, in the chorus it changes its timbre, more attack and more presence, and on the bridge, it is wider and more sustained, making the whole section different.
- Using distortion to make timbrical changes in the vocal in different sections.
- Leaving the highest end of the spectrum always free for the air of the lead vocal. The background vocals are equalized and placed in the stereo field to avoid masking the lead.
- Reverb density affects the timbral characteristics of the vocals. We can contrast lead vocals from background vocals by using different reverbs, which will add different harmonics and therefore bring differentiation to the center of the stereo field.
- Using compression dynamically. Controlling the loudness range so towards the end of the song there are less dynamics and an increased energy feel.

4.3 Analysis 3: Vulnerable - Selena Gomez

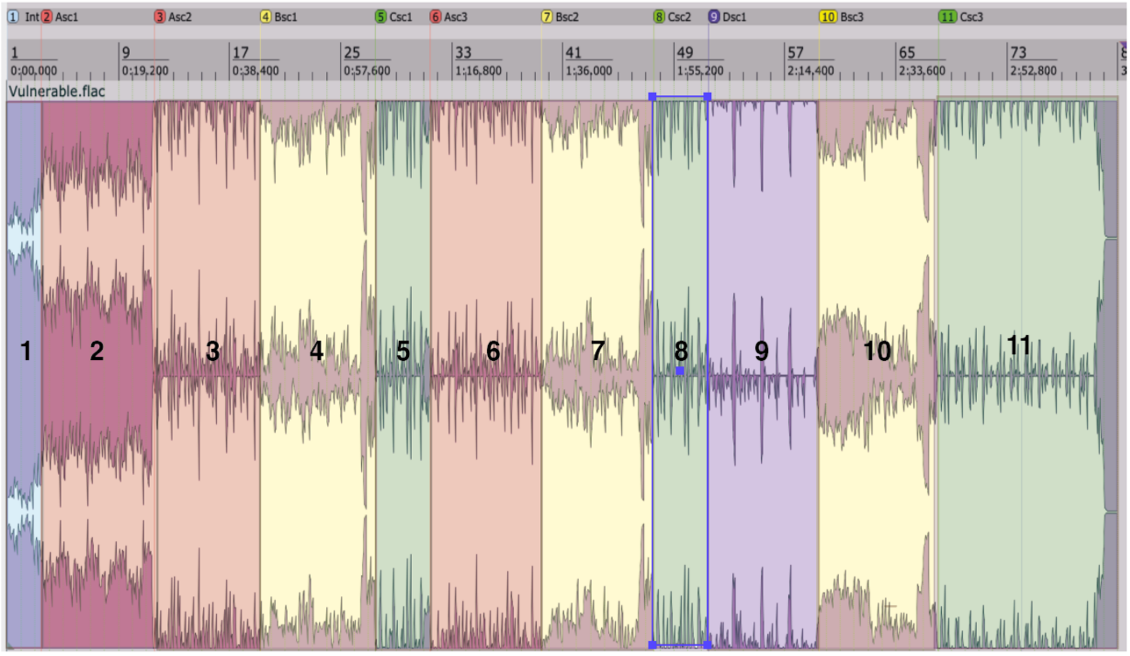

Figure 4.3.1: *Structure of "Vulnerable" by Selena Gomez.*

4.3.1 Presentation and musical acknowledgements:

The producers of the song are: The Monsters and the Strangerz, Jon Bellion, Schoudel and Gian Stone. The Monsters and the Strangerz are a group of 7 songwriters and producers based in LA who have worked with major pop artists like Katy Perry, Miley Cyrus, Dua Lipa, etc. and whom Jon Bellion started working with Jon Bellion in 2018, as coproducers; they are now frequent collaborators. Bart Schoudel is also a top-notch LA producer, who has also worked with big artist such as Beyonce, Justin Bieber, Camila Cabello, etc. He is the only producer who has credits in more than one song of the album. Lastly, Gian Stone is a songwriter and producer from New York who also worked on Dua Lipa's Physical, and was number one on the Billboard Hot 100 Producer Chart in May 2020 for his work with Ariana Grande and Justin Bieber.

The song is in the key of Bb major and the tempo is 100 bpm. It is a danceable song but intimate at the same time, as the lyrics talk about insecurities in a relationship. It has minimalistic production and a very clean sound. The vocal is the element that stands out the most. The performance is calm, it leaves long spaces in between phrases that get filled with reverbs, delays and spacey background vocals, that give an atmospheric vibe. The instrumental bridge also adds to this ethereal feeling and it is a very surprising section.

The song was mixed by Serban Ghenea and mastered by Chris Gehringer, the all-star sound engineering duo of today's pop. They are based at Sterling Sound studios in LA and have worked on the biggest selling pop albums of the last decade. Serban Ghenea is the mixer for most of Selena's songs, unfortunately he is a very secretive figure, who never gives interviews or reveals his techniques.

4.3.2 Structure

1	0:06	Intro	An automated low-pass filter at approx. 600Hz starts the song. Instrumentation inside this filter consists in: Kick drum, a wide pad laying the harmonic foundation and a rhythmic bass synth that does a crescendo.
2	0:25	A sc1	The filter opens to around 1kHz. Percussion: filtered between 100-1kHz. Electronic sounds of kick, snare and shaker and an occasional cowbell type of sound. Bass: 150-500Hz not mono. Pad: same as Intro. Vocals: It is bright, high-mids and air (>10kHz). Dip around 1-2kHz. Stereo short reverb with slapback. There's flanger on the reverb that can be clearly heard listening to the sides. The reverb doesn't decay, it gets drastically cut at the end of every sentence.
3	0:44	A sc2	The filter is opened for all the instrumentation.

			Kick drum “four on the floor” beat. Fundamental at 100Hz, very round, doesn’t have a lot of attack. Snare is high pitched (+-5k), and in the middle. There are percussive elements on stereo, filtered between 300-1kHz, subdividing the beat and accenting 2 and 4. High pitched shaker (4-10kHz) very wide and to the sides, also accenting 2 and 4. Bass: Very subby (fundamental at 40Hz) and sustained. It follows the harmonic rhythm of the synth. It is very compressed; every note sounds the same dynamically. The mids are stereo (>200Hz). Guitar: funky style, subdividing in 16^{th} notes and accenting the upbeats. It is played with a pick and we can hear the attack more than the pitch. It works more as a percussive element than a melodic one. Synths: - Intro pad + louder pad, wide in the stereo. It swells in in every chord like the bass, it is probably a little bit sidechained. - Melodic pad (0:33’) boosted at 1kHz. Sits on top of harmonic pad and very stereo. Vocals: Deeper dip around 1-2kHz than Asc1. No flanger in the reverb. There’s two vocal lines. The one which is higher in pitch takes the first plane. Every 2 bars there’s a vocal delay with lots of feedBack, high passed and with reverb that sends it to the back. It decays in volume and is pan automated to the left the first time and to the right the second time, disappearing before the new lead sentence. As a transition there’s a riser that is probably the pad or a vocal reversed and with reverb, and at the same time a descending pitch line contrary to it.
4	1:03	B sc1	Kick drum: very filtered “bomb”- like sound (40-100Hz) with big reverb. Perc: Filtered percussion crescendo to transition to chorus.

			Bass: Very wide (it has some slow (quarter note?) tremolo effect?). From the 5th bar (0:55') it is layered with another sound, that adds attack and distortion. Synth: Pad does crescendo. Filtered 300-1kHz. Ends with reverse effect. Guitar disappears in this part. Vocals: The vocal is doubled and panned L-R. Overall brighter timbre than before. Different reverb: large, bright and wide. We can hear slapback and it is louder. At the end big dip at 2kHz, lot of 10kHz air. Same slapback as Asc1.
5	1:13	C sc1	Perc: Same as Asc2. Bass: Same timbre as Asc2 but louder. Guitar comes in again, same timbre as in Asc2. Synth: Same pad + vocal chops pan automated to L- R + melodic details with big reverbs sent to the back and pan automated. Vocal: More body (energy around 200Hz). Same FXs as Bsc1, but at the end of every "I'll stay vulnerable" the decay of the reverb is cut. Background vocals: 2 planes. 1st- "yeah", loud and same timbre as lead, 2nd is far, sent to the back with bright reflective reverb. Reverb tone emphasizes 500Hz.
6	1:32	A sc3	Same as Asc2 + new background vocal harmony: higher, not doubled + other background responses to main vocal, darker and in the middle but sent to a different reverb, further, in a different space.
7	1:51	B sc2	Bass, drums and synth same as Bsc1. Same vocal timbre and effects (maybe automated decay rise at 1:50') as Bsc1 + high pass filtered harmony vocals, with emphasis on 2k. The high end is left only for the lead vocal.

			At the end of the section, we can hear the reverb is sort of harmonized, maybe sent to an auto-tune effect.
8	2:01	C sc2	Same as Csc1. Transition with reverse swell.
9	2:20	D sc1	Perc: The basic beat, kick, snare and shaker, are the same sounds but now stereo toms come in. 2:10' tambourine enters and adds a crescendo feel. Also, there is a noisy, "wind" stereo sound that works as a pad. The percussion is louder here than in all the other parts. Bass: stereo and softer, low end not as perceivable. Synth: Pad + vocal chops automated in panning. Ambient feel. Different planes of vocal chops.
10	2:40	B sc3	Bass: enters 2:25' and does a crescendo. Same timbre as the B's. Synth: first pad then crescendo adding all previous parts. Vocals: New sound, bathed in reverb of long decay. The reverb timbre emphasizes 200Hz and 6kHz. It gets drastically cut at 2:39.
11	3:12	C sc3	Same instrumentation as Csc2 + pan automated L-R Synth 3:00'. Guitar does the same pattern but doing high pitched octaves (sounds more present). There are added background vocal harmonies and ad libs in the middle. Final chorus has even more harmonized vocals and the main melody is a bit changed. The sound contrast with the end, the last "I'll stay vulnerable" is a low, dry vocal which has the same timbre and effects as Asc2.

4.3.3 Frequency/Timbre

Vocals and reverbs are the main characters of this song. The effects on the first verse of the song are a statement. Selena is placed in an unreal space, due to the "flangery" sound and the the way the effects get cut off without decay, we get a feeling of uncertainty. Her vocal is warm and feels close but bright. As we can see in Fig 4.3.2 the shape of the spectrum resembles a smiley face. The emphasized air makes us hear details of mouth noises, which have not been removed, to achieve intimacy. The reverbs and delays throughout the song also feel un-real. They have been equalized to reduce mid-low energy. We can hear it at 1:02, for example, where we can't hear energy below 300Hz or 1:50', where the reverb is different but Eq'd high-passed in the same way.
Bsc3 is the only place where the reverb decay of the lead vocal is not cut off. As we can see in Fig 4.3.3 the reverb fills the mids that the vocal lacks of, and decays on the high end, to keep the vocal close and detailed.

Figure 4.3.2: *Bsc3 Vocals mono spectrum.*

Figure 4.3.3: *Bsc3 Vocals side spectrum.*

In fig 4.3.4 we are able to see the overall spectrum of the song in the most crowded section, the last chorus. It is fairly flat. We can identify a dip around 150Hz, which is where some boominess can usually be found, and little dips at strident or sibilant frequencies on the mid-highs.
Serban Ghenea was also the mixer on the song Lose You to Love Me, from the same album, which has a similar ambient and vocal centered vibe. If we look at the spectrum of the last chorus of this song (Fig 4.3.5) we can see similarities with Vulnerable, like the dip around 150Hz and similar low end and midrange, although "Vulnerable" is considerably brighter.

Figure 4.3.4: *Csc3 average song spectrum.*

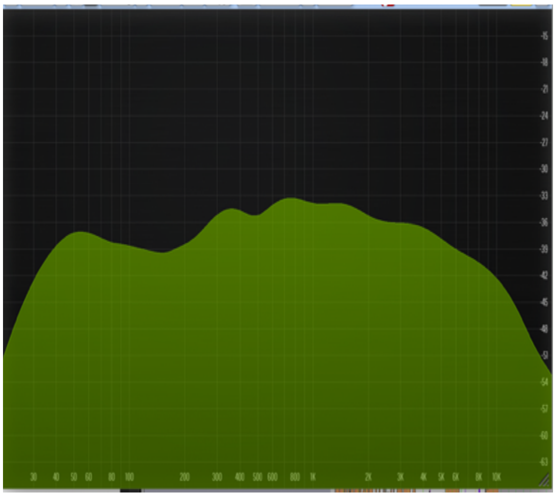

Figure 4.3.5: *Last chorus average spectrum on" Lose You to Love Me".*

4.3.4 Space

The start of the song, the treatment of the vocals, the “flangery”-reverb that doesn’t decay, tells us that the space we are in is somewhat unhuman. If we listen further, we realize that the depth of the whole image is very big. The vocal chops and ad libs all have big reverbs, making us feel like the virtual space we are placed in is a very large empty space. At the same time, the brightness of these reverbs is always present and very audible, which is a phenomenon that can never be experienced in real life, because of distance losses, therefore, the space where the song is set is a physically impossible place.
In figures 4.3.6 and 4.3.7 we can see how the song widens after the bridge.

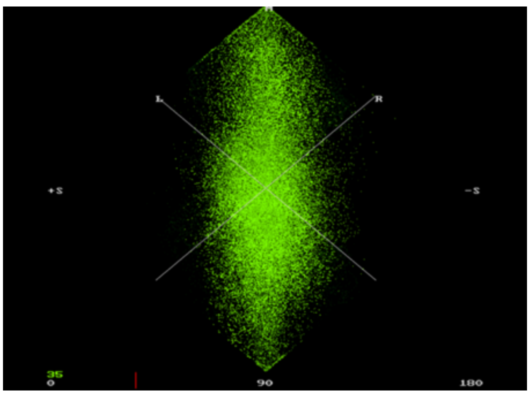

Figure 4.3.6: *Chorus Csc2*. Last chorus before the bridge.

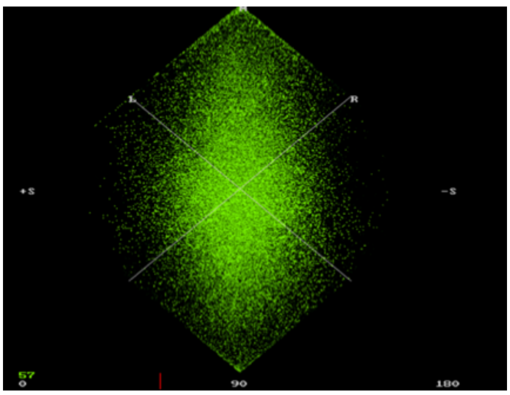

Figure 4.3.7: *Chorus Csc3*. First chorus after the bridge.

It is interesting in this case to relate the song's space to its lyrics. The song talks about vulnerability and reveals some deep thoughts from the singer, which would typically be translated into an intimate production, but in contrast, the song is massive. If we take away this element of space and listen to the song in mono, the point of the song changes completely. Although it has been masterfully treated and every element is still audible (we don't lose any melodic or harmonic content when we sum to mono) the central figures of the song change. The beat and bass take the spotlight and it feels more like a groovy club pop song. If we listen to the whole song once in mono and then in stereo, we can perceive again, as if we were listening from the first time, the importance of the reverbs and effects this song has, and how they relate to its purpose. This song doesn't make sense without the space, because it is what gives us the sensation of opening up, the song is liberated from an enclosed space, which is the musical translation of the feeling of letting go and liberating from past experiences, which is what ultimately the lyrics are about.

On an interview with Waves audio, Serban Ghenea spoke about his approach to mixing and said: "For any given song, you need to listen and figure what it is in the song that affects you, what it is that touches you – and then try to maximize that. You need to find whatever is the special thing about that song."[43] In this case, the reverb treatments could be that "special thing".

4.3.5 Loudness/Dynamics

This is the most dynamic song we have seen so far. We can see it only by looking at the waveform (Fig. 4.3.11), there are peaks and valleys and we can tell that dynamics play a role in the structure. It has lots of moments of silence, only reverbs, only vocals, etc. that contrast with the most crowded parts. Maintaining the dynamics throughout the song helps with making the song a progressive crescendo and helps make climactic the last chorus.

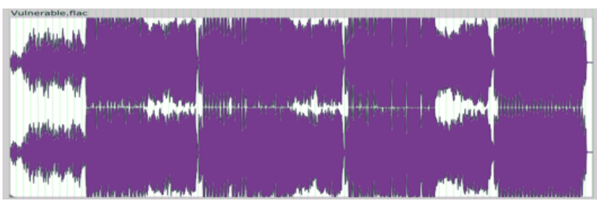

Figure 4.3.11*: "Vulnerable" Waveform.*

[43] Waves. (2021). Waves x Grammy 2021: Serban Ghenea Mixing Taylor Swift. 14 abril 2022, Waves Audio. https://www.waves.com/waves-grammys-2021-serban-ghenea-taylor-swift

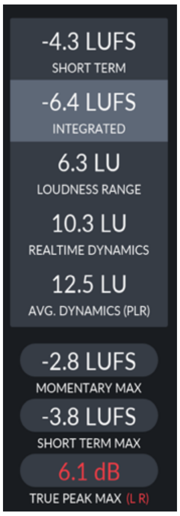

Figure 4.3.9: *Loudness measurements of the whole song. on Youlean Loudness Meter*

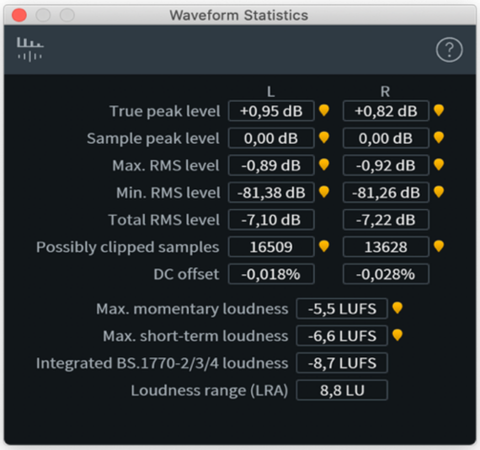

Figure 4.3.8: *Loudnes measurements for the whole song on Izotope RX8 Waveform Statistics.*

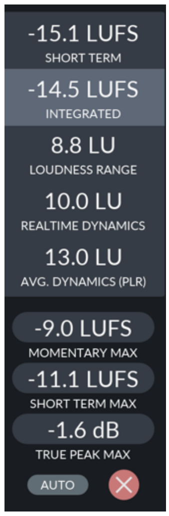

Figure 4.3.10: *Loudness measurements of vocal dynamics.*

4.3.6 Conclusive thoughts

As we have seen, the production and mixing have a key role in bringing out the meaning of a song. The space created takes a big role in this case, and the use of effects in a subtle way is key. Clearly the vocal is the main element and we have seen there are different vocal scenes throughout the sections. None of them are made too obvious. The transitions are subtle but they are what keeps us interested and expecting. All in all, the processing helps to tell the story, enhancing the main element, the vocal performance, in a delicate manner and adding a layer of meaning.

Some techniques learnt:

- Leaving vocal air only for the lead vocal and keeping the low instruments in the low-frequency register, to avoid cluttering the voice.
- Reverb equalization to fill spaces where vocals don't have much energy.
- Including sonic "details" in the sections' transitional silences (e.g. a reverb tail that is "tuned", a glissando filling the silence).
- Using filter automations and crescendos to bring dynamism and make transitions between sections.
- Use of reverbs to create different virtual spaces. Manipulate them to create impossible scenarios or cut them suddenly to bring interest.
- Leaving reverb decay only for background vocals so that decay does not blur the lead vocal.

4.4 Analysis 4: Rager teenager! - Troye Sivan

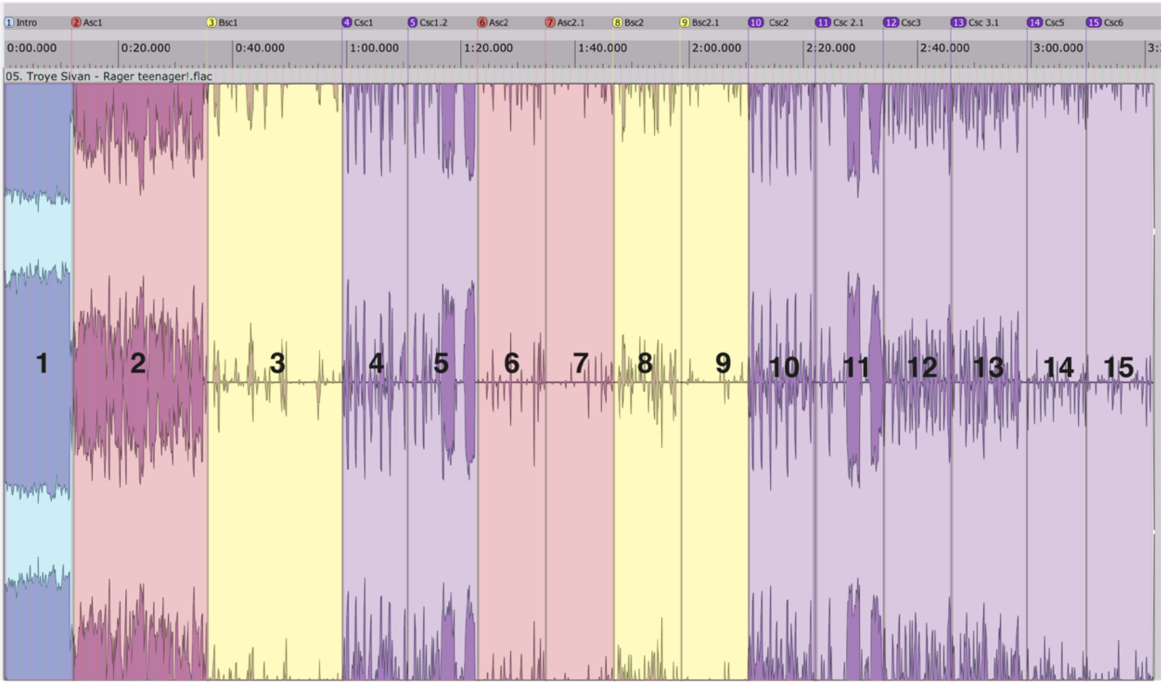

Figure 4.4.1: *Structure of "Rager teenager!" by Troye Sivan.*

4.4.1 Presentation and musical acknowledgements:

The song talks about teenagerhood and has an electro-pop vibe, although the tempo is slow, 81bpm, and it is quite minimalistic. It is in the key of D major and the main motif is a harmonic descending chromatic line the synths make throughout the song. Curiously, the chorus is the simplest part of the song; although the verse creates expectation and builds up, eventually it falls when the chorus arrives, in this way creating movement and unpredictability. The main elements are the vocals, which are displayed very wide in the frequency and stereo spectrum.

The song is mixed and mastered by the previously mentioned duo, Serban Ghenea and Chris Gheringer. The mix is wide and full of contrast, and the whole album, all mastered by Gheringer, is remarkably cohesive although its very different songs.

4.4.2 Structure

1	0:12	Intro	Synth pad. Narrow spectrum (200Hz-1.5kHz). Stereo but not very wide. It has a 1/8th note auto-pan subtle effect.
2	0:35	A sc1	Same synth. Lead vocal: It is upfront and with body. It has a mid-sized reverb (with pre-delay) that is very reflective (we can hear it at 0:23). Background vocals: Wide in stereo. They are band-pass filtered; they resonate on the midrange (around 1-2kHz). They have distortion which makes them sound granular and are sent to a small room reverb. They accentuate the first beat of each bar.
3	0:59	B sc1	Bass: Distorted and stereo. It also has a subtle auto-pan effect. It is very compressed and has lots of sustain. The attack is not prominent. The sub-bass is (strongly) emphasized but also mid-range from the distortion. More synths: They widen the stereo image through the midrange. - String synths higher octave. - Moog-like filtered synth (brassy). - 0:53 more aggressive synth sound does a crescendo. Vocals: The lead vocal is doubled in L-R and sent to a bigger reverb now without pre-delay, which sinks it more into the space. The reverb decays before the next sentence, almost as if it was automated (we can hear it listening to the vocal stem and the sides only). The timbre of the vocals as a whole is brighter, it has a lot of air, and fills more of the frequency spectrum. It also has a filtered stereo delay. Overall, they sound a bit grainy, and with texture, which is probably the effect of added distortion or tape emulation. The background vocals are in unison with the lead and very wide and bright.

4	1:10	C sc1	Synth: same motif as the intro but the timbre is now a sustained rough pad with less attack. It is also placed wider in the stereo. Vocals: Layers of unisons. In the middle, the lead vocal is doubled and also harmonized an octave higher. In the sides we can hear the lead melody and the octave also playing, which means we have a stack of (at least) 5 vocals. The lead one has more body and warmth than the rest, which helps differentiate it. They are all sent to a short reverb from which we don't hear any slapback. This reverb is resonant at around 250Hz, we can hear it at 1:04 on the word *out*. The high end in this vocal is very emphasized and it becomes very sibilant and bit strident.
5	1:22	Csc1.2	A very subtle harmony vocal comes in and the reverb becomes like a shimmer type of reverb.
6	1:34	A sc2	Beat: The sounds are not of a natural drum set. The low end is occupied by a very low frequency kick drum with a big reverb that is sustained throughout the loop, and a tom, that is more active and sits around 80Hz. There are several clap sounds. In the middle they create a looped pattern, and on the sides, others add details and open up the beat. The sound of the beat is distorted and broken. It sounds as if they lowered the bit depth and sample rate of the file to achieve this kind of distortion and low-resolution. Bass: the bass is very subtle. Same timbre as Bsc1 but lower in volume. Synth: Same as Asc1. Vocal and Backgrounds: Same as Asc1.
7	1:46	A sc2.1	More vocal harmonies are added in the stereo. They are doubled and band-pass filtered at around 2kHz. The sound reminds of the sound of a choir (1:39').

8	2:10	B sc2	Beat: only kick drum. Different sound. Very filtered and resonant at 60Hz (reminds of a heartbeat). Bass: Same sound as Bsc1 but a bit lower in volume. Synth: Same instrumentation as Bsc1 but initially lower in volume. It does a crescendo to go to Bsc2.1. Lead vocal: Same as Bsc1. Background vocals: They have a ping pong delay, set at quarter note rhythm and very distorted and occupy 1kHz-5kHz spectrally.
9	2:10	B sc2.1	Synth: The synths have more presence than the strings. Here they have more aggressive timbres and the auto-pan and effects create a phaser sensation. Electric guitar: there is one doing very distorted chords (that could be a synth) in stereo and one that is closer, doing a melody panned to the right, which adds variation. The main idea behind this part is the crescendo towards Csc2.
10	2:22	C sc2	Beat: snare stick, it is in the middle but has a moderately big plate reverb in stereo that places it behind the vocal and widens it. No bass. Synth: Same synth as Csc1 + a sound similar to vocal chops or synthesized vocals plays every quarter note, in the background. Vocals: Same as Csc1.
11	2:34	C sc2.1	Kick drum comes in, same filtered sound as Bsc2. Snare remains the same. Background vocal harmonies. Vocals: Same as Csc1.2 At the end of the section a harmonized melody of two distorted synths does a crescendo, one panned to the left and one to the right. They fall into the melody of the next section.
12	2:45	C sc3	Beat: Kick drum sound is the filtered one and plays every beat. There is no tom. The snare sound has a multi-tap delay which is

			automated in panning (goes from hard right to hard left) and in frequency (a low-pass filter goes lower in each repetition) every two bars. The sound of the snare seems to be filtered white noise. The general timbre is still distorted but the whole drums have room reverb and are not as aggressive. Bass: Same as Asc2.1. Synth: very shrill timbre. It is more on the sides than in the middle. It has a long reverb that doesn't decay, almost like a freezed reverb sound. It also has a short delay, that adds depth to the sound.
13	2:59	C sc3.1	More synths come in: A new pad, a glissando line with stereo delay add depth to the mix. Finally, an arpeggiated filtered synth comes in towards the end of the part, as well as high-pass filtered noise, which appears at 2:57' filling the 4kHz band.
14	3:09	Csc3.2	The arpeggiated synth grows in volume and the filter gradually opens.White noise raisers are added in stereo to create a buildup.
15	3:21	Csc3.3	The arpeggiated synth is totally opened and goes to the higher octave. The noises are increased in volume creating a chaotic soundscape, that is the climactic moment of the song. Everything ends suddenly, taken away by an automated low-pass filter.

4.4.3 Frequency/Timbre

There are many spectral differences between the voice on Asc1 (Fig 4.4.2) vs Bsc1 (4.4.3). On the first one, the lead vocal has a lot of prominence around 400Hz, making it feel close and deep, as if he sang very close to the microphone. The background vocals sound boxy (between 300Hz and 2kHz) and distorted, I assume with the intention of making them complementary from the lead, probably because they wanted to keep the verse narrow, so that there could be a stereo explosion on Bsc1 (Fig 4.4.5). On the next part, Bsc1, the lead has a lot more brilliance and less body than before, and the backgrounds are very similarly processed, making it hard to differentiate the lead from the rest. In the next part, C sc1, (Fig 4.4.3), the vocal timbre is a mix between the first part and the second. Again, we have a lead vocal that sounds close and deep, with an even more prominent bump around 400Hz but now with more high end too. The backgrounds also have a lot of high end and are at a similar level to the lead, altogether forming an ensemble of singers placed in a big lush reverb.

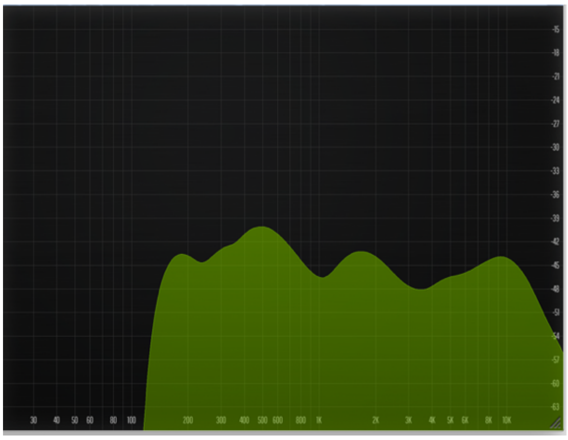

Figure 4.4.2: *Vocal spectrum on Asc1.*

Figure 4.4.3: *Vocal spectrum on Bsc1.*

Figure 4.4.4: *Vocal spectrum on Csc1.*

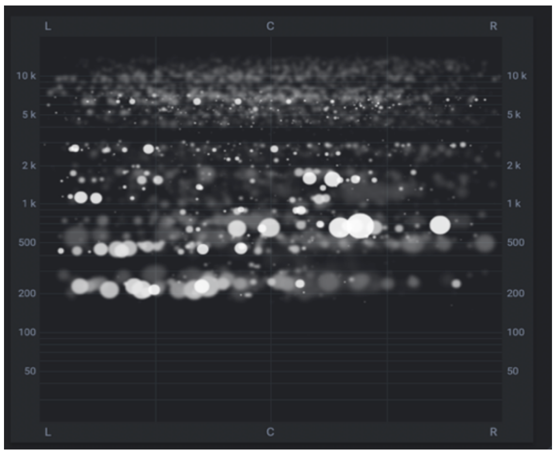

Figure 4.4.5: *Vocal stereo and frequency spread on Bsc1.*

In Csc1, the hook, it is the contrast and lack of meeting the expectations that creates a unique moment. Where the low end is not used and there are voices and the harmony synth, which is narrow in frequency creating a very minimalistic moment where the vocals shine. In figures *4.4.6 and 4.4.7*, we can see the space left in the music stem for the vocals to occupy.

Figure 4.4.6: *Music spectrum on Bsc1.*

Figure 4.4.72: *Music spectrum on Csc1.*

In conclusion, we can see that the treatment of the vocals is constantly seeking for contrast and surprise, and that in some moments the backgrounds and the lead merge into one, creating a beyond-human voice, maybe an inner voice that speaks about this person's desires, whereas the other one is the more realistic.

4.4.4 Space

In terms of space we can divide the song in two big sections. In the first half of the song (until 1:22) there is no beat or mono bass, and during the first Bs and Cs, the sense of the middle is not clear and overall the sound is wide and spacious. On the second half, the beat comes in, bringing a grounding feeling and, although the stereo is equally wide on the music and vocals, it makes everything more centered and stable. We can see the progression of the stereo width in Fig 4.4.8. Each section has a different width, which grows as the song progresses.

The same contrasts we have found in the timbrical treatment of the vocals, are very aligned to the spatial changes. When the vocal has more low-end it is placed closer and, in the middle, whereas when it gets brighter it is placed wide in the stereo and further in the mix. This contrast makes the song very interesting and it captures our attention. It is also a counter-intuitive phenomenon, because in real life, when a source gets further away, we start to lose high end definition, not the opposite.

Another element that adds interest to the song is movement. As stated in the structure, the main harmonic synth and the bass have an auto-panning effect, that is always on, adding subtle movement behind the vocals. Although it is not an exaggerated effect, I believe we perceive it unconsciously, in the way it makes us move while we listen. The snare sound that begins in C sc3 also has movement, this time in space, from left to right, and in frequency, because it has an automated filter. These effects, make the song less static and when there's minimal instrumentation, like in Asc1, it provides a level more of intricacy in a seemingly simple mix.

Lastly, if we listen to the song in mono, apart from losing the sense of space and ambience, we also lose a lot of the aggressiveness, because the high-end is very emphasized and distortion from the drums and the synths in the sides, which create the very particular shrill sound of this song, which in mono gets lost.

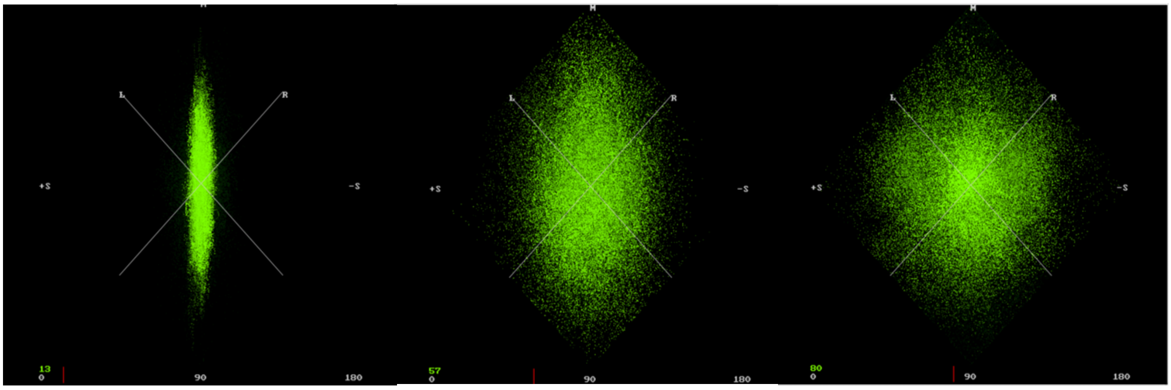

Figure 4.4.8: *Progression of stereo wideness from Asc1 to Bsc1 to Csc1.*

4.4.5 Loudness/Dynamics

In figures 4.4.9 and 4.4.10, we can see the integrated loudness measure is 7,7 LUFS and the true peak exceeds 0dBs. I checked it twice with two different audio files, to be sure that it was not a file error, and both measurements turned out the same. This track is mastered really loud (even louder than the Selena Gomez track analyzed earlier, which was mastered as well by Chris Gheringer). The Loudness Range, is above 5LUs, which means there are still some dynamics preserved.

Figure 4.4.9: *Rager teenager! loudness measurements on Izotope RX Waveform Statistics.*

Figure 4.4.10*: Rager teenager! loudness measurements on Youlean Loudness Meter.*

In figure 4.4.11 and 4.4.12, we see that the Loudness Range before the beat comes in is 11,3 LU, because the first two sections, intro and Asc1, are softer (we can see it only by looking at the waveform). In the second half, when the beat starts, and until the end, the Loudness Range is 2,9 LU. Still, in the two sections the integrated loudness measurement is not very different (difference of 1 LUFS), which shows that the song is really compressed and limited.

Figure 4.4.11: *Loudness measurements before the beat comes in (0:00'-1:23').*

Figure 4.4.12: *Loudness measurements after the beat comes in until the end (1;23'-3:20').*

4.4.6 Conclusive thoughts

It feels like this mix is trying to externalize the polarized emotions of being a young boy. The mix presents us his dreams and hopes as big spaces and lush bright vocals, that sound unreal and imaginative, and, on the other hand, we have the real-life situation, more intimate, narrow, and a little hopeless. The drums and backgrounds give us the anger of this rager teenager whereas the choruses sound more like his desires. It feels that thanks to the soundscapes created, we can get inside the mind of this person, and feel his own contradictions.

Some techniques learnt:

- Use of auto-pan effect to create almost imperceptible movement in rhythmically un-active moments.
- Automating equalization filters.
- Awareness of the placement of the distortion. If this is placed in the sides, part of the aggressiveness will be lost when in mono.
- Change timbre and spatial characteristics in parallel, to exaggerate meaning and emphasize contrast.

5. Conclusions

The present study has proposed and put into practice a methodology for the sonic analysis of contemporary pop music. In order to achieve this, first, I studied the sonic characteristics, post-production practices and figures involved in the creation of this music, in order to understand its sonic footprint. With this information, and mainly based upon the model of William Moylan's "Recording Analysis: How the record shapes the song" (Moylan, 2020), I constructed my own methodology, with the objective of extracting the essentials of Moylan's teachings (which I found to be very enlightening but very demanding and hard to put in practice), and, at the same time, fulfill some of the informative gap in the field of critical listening. I hope the work to serve as a guide to understanding the needs, requirements, and processes that a sonic analysis must take into account, providing, both, to amateurs and to sound professionals, with a pathway to accomplish rigorous analysis. Subsequently, as a practical application, I worked out and presented four analyses of hit songs of 2020. At the end of each analysis, I constructed a list of the most interesting production and mixing practices I found. From them, I have been able to derive some commonalities between the four tracks:

- Achieve separation between instruments through the use of complementary equalization (i.e. separating lead and background vocals to avoid masking), and spatial difference (i.e. widening the bass in order to leave the center for the kick drum or using different reverbs to differentiate spectrally similar elements).
- Use of FX to achieve timbral contrast (i.e. different reverb types and densities or use of distortion).
- Achieve the sensation of dynamics through the arrangement while maintaining a loud average volume. The arrangements include minimalistic sections, which contrast the fullness of the climactic parts, in this way creating the sensation of dynamics, although maintaining an overall compressed sound.
- Increasing sub-division of percussive loops and melodic detailing as the song progresses to avoid stillness. The subdivisions and melodic details add layers of depth and keep the ear interested throughout the song.
- Use of automation of panning, filters, reverb, delays, etc. to create interest and movement.

To support the analysis, I constructed a graphical representation method, which, because of the lack of time, is only presented in the form of a sample graph in section 3.5, with the purpose to apply it in future analyses.

I believe I have been able to conform a systematic approach to sonic analysis, by giving insight into what to listen for and what to visualize, in order to understand and describe the key processes in making pop music sound the way it does. I would like to encourage any reader to listen to one of the songs first without reading my analysis, and then have one or two listens while reading the text. I believe this could make anyone grow their appreciation and interest for the presented music and be able to see the layers of meaning the production and post-production process convey.

The main challenge encountered during the creation of this methodology, and one I am not certain I have overcome, has been enclosing the system inside a musical genre. I presumed, at the start of this journey, that, by reviewing the existing methodologies I could infuse my findings into the genre I was interested in analyzing, achieving, in turn, a very focused method. Along the way I have found that the key point to achieve this does not rely on the analysis tools, but on the analysis approach, and that still, with this in mind, a lot of different genres are treated in a similar way in post-production, which ultimately requires, the analysis approach to shift from song to song, in order to better describe each one. In the end, I resolved that the approach for pop music analysis should be focused around the sonority of the voice and the rhythmic components, considering that they are the key elements, which carry the essence of its sonority. Ultimately, I believe this approach can also be valid for other genres, making me learn that no one methodology can be genre-specific, but that it can be adapted to each type of musical setting.

Another challenge has been deciding to what extent to rely on the information given by visualization tools. Music is a very subjective phenomenon we try to make objective through structured processes and graphs, and it undoubtedly can help us reach to conclusions that our ears cannot alone reach. At the same time, it sometimes can make us observe things we cannot hear or viceversa (e.g. when we don't have enough resolution to see low end information peaks) and in this way the analysis can be skewed. Information on the set up of the tools is essential to avoid this, but I believe we should always contrast the information we get from analysis tools with our perception, and we should avoid talking about the elements we see on a graph but that we cannot perceive. It is through practice, that we can improve our critical listening skills, and progressively be able to grasp details and describe sonorities.

The culmination of this work ends in its use. This method is susceptible of being furtherly refined and improved through its use and test with different sources. In the future, work can be expanded towards the automatization of the graphic representations presented, allowing to obtain visuals for any song and help assimilate and transmit the analysis findings. The method can also be expanded in the description of effects and their role, furtherly exploring why certain effects are chosen and the contribution they make to the significance of the song. Another idea that could not be fulfilled due to lack of time, but I hope to develop in the future, is the creation of educational videos, in the style of Youtuber Adam Neely, for example, to transform the written analyses into dynamic videos that can appeal to anyone interested in sonic analysis and help spread knowledge on the matter of critical listening.

In addition to the lessons learned on the matter of analysis, this work has also served as great insight into the work and engineering approach of a few remarkable sound engineers. From my practical analyses I have been able to derive some of the techniques used to create contrast, definition, balance, etc. with the objective of developing and reinforcing the musical narrative. I will undoubtedly incorporate these findings in future work and hopefully the present "Treball final de grau" (Bachelor's thesis) can be of use to other people in the sound engineering field.

References

- Askin, N & Mauskapf, M. (2017). *What Makes Popular Culture Popular? Product Features and Optimal Differentiation in Music*. American Sociological Review, 82 (5), 910-944. 20 April 2022, SAGE Database.
- Tolsen. (2022, May 4). Billboard hot 100. Billboard. Retrieved May 14, 2022, from https://www.billboard.com/charts/hot-100/1950-01-07/
- Blacking, J. (1977). *Some Problems of Theory and Method in the Study of Musical Change*. Yearbook of the International Folk Music Council Vol. 9 (1-26). New York: Cambridge University Press.
- Burgess, R. (2013). *The Art of Music Production*. New York: Oxford University Press.
- Bruce, D. [David Bruce Composer]. May 18, 2020. *The Subtle Art of Modern Pop Production*. [Video] Url: https://www.youtube.com/watch?v=dwexkS4Djb8
- Cook, N. Clarke, E. Leech-Wilkinson, D. & Rink. J. (2009). *The Cambridge Companion to Recorded Music*. New York: Cambridge University Press.
- Corey, J & Benson, D.H. (2016). *Audio Production and Critical Listening: Technical Ear Training*. New York: Routledge.
- Dockwray, R & Moore, A. (2010). *Configuring the sound-box 1965–1972*. Popular Music, 29 Issue 2, 181-197.
- Eno, B. (1983). *The Studio as a Compositional Tool*. Down Beat magazine, July 1983. p.56.
- Fink, R. Latour, M. & Wallmark, Z. (2018). *The Relentless Pursuit of Tone: Timbre in Popular Music*. New York: Oxford University Press.
- Frampton, T. (2019). *How Pros Make Hits*. London: Mastering the Mix LTD.
- Frith, S. Straw, W. & Street, J. (2001*). The Cambridge Companion to Pop and Rock*. New York: Cambridge University Press.
- Frith, S & Zagorski-Thomas, S. (2012). *The Art of Record Production*. England: Ashgate.
- Gibson, D. (1997). *The Art of Mixing: A visual Guide to Recording Engineering and Production*. California: Mix Books.
- ITU-R. (March 2017). *Methods for selecting and describing attributes and terms, in the preparation of subjective tests.* Report. Geneva, Switzerland: International Telecommunication Union Radiocommunication Assembly (ITU-R).

- Jacoby, S. [Waves Audio]. October 21, 2019. *1 Song, 2 Mixes. How Mixing Can Transform a Song.* Url: https://www.youtube.com/watch?v=RJX-V1F8j0w&t=204s
- Leviatan, Y. (2022). *Making Music: The 6 Stages of Music Production*. March 22, 2022. Waves Audio Url: https://www.waves.com/six-stages-of-music-production
- Martin, P.J. (1996). *Sound and society: themes in the sociology of music*. Manchester: Manchester University Press.
- Mastering the Mix. (2021). *What Are Dynamics in Music?* 9 April, 2022. Mastering the Mix. Url: https://www.masteringthemix.com/blogs/learn/what-are-dynamics-in-music
- McAdams, S. (1999). *Perspectives on the Contribution of Timbre to Musical Structure*. Computer Music Journal 23. (3).
- Middleton, R. Buckley, D. Walser, R. Laing, D & Manuel, P. (2001). *Pop.* January 20, 2022. The New Grove Dictionary of Music and Musicians Url: https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000046845
- Moore, A. (2003). *Analyzing Popular Music.* New York: Cambridge University Press.
- Moore, A. *Song Means: Analysing and Interpreting Recorded Popular Song.* England: Ashgate.
- Moylan, W. (2020). *Recording Analysis: How the Record Shapes the Song*. New York: Routledge.
- Moylan, W. (2007). *Understanding and Crafting the Mix*: The Art of Recording. Oxford: Focal Press.
- Myers, K. (2016). *The History of Pop Music In 5 Defining Decades*. 5 December, 2021, Culture Trip. Url: https://theculturetrip.com/north-america/usa/california/articles/the-history-of-pop-music-in-5-defining-decades/
- National Instruments Corp. (2022). *Understanding FFTs and Windowing*. 18 April, 2022. National Instruments Corp. Url: https://download.ni.com/evaluation/pxi/Understanding%20FFTs%20and%20Windowing.pdf
- Ni, Y. Santos-Rodríguez, R. McVicar, M. & De Bie, T. (2011). *Hit Song Science Once Again a Science?* University of Bristol. Retreived from: http://www.tijldebie.net/publications/scoreahit_paper

- Pedersen, T. (January 2008). *The Semantic Space of Sounds: Lexicon of Sound-Describing Words*. Denmark: DELTA.
- Rauscher, C. Volker & J. Roland, M. (2001) *Fundamentals of Spectrum Analysis*. München: Rohde & Schwarz.
- Schmidt Horning, S. (2012). *The Sounds of Space: Studio as Instrument in the Era of High Fidelity*. The Art of Record Production (30). New York: Ashgate.
- Serrà J, Corral A, Boguñá M, Haro M, Arcos JL. Measuring the evolution of contemporary western popular music. Scientific reports. 2012;(2):521. DOI: 10.1038/srep00521
- I. Shepherd, E. Grimm, P. Tapper, M. Kahsnitz, and I. Kerr, *Measuring Micro-Dynamics—A First Step: Standardizing PSR, the Peak to Short-Term Loudness Ratio,* Engineering Brief 373, (2017 October.). Retrieved from: https://www.aes.org/e-lib/browse.cfm?elib=19324
- Shepherd, I. (2022). *What is the Loudness War?* 28 March, 2022. Mastering Media Ltd. Url: https://dynamicrangeday.co.uk/about/
- Sound on Sound. *Glossary*. February 9, 2022. 17-04-22, de Sound on Sound. Url: https://www.soundonsound.com/glossary
- TC Electronic. (2022). *Loudness Explained*. April 10, 2022, de TC Electronic. Url: https://www.tcelectronic.com/loudness-explained.html
- Théberg, P. (2001). *Plugged in: technology and popular music*. The Cambridge Companion to Pop and Rock (3-24). New York: Cambridge University Press.
- Theberge, P. (1989). *The 'Sound' of Music: technological rationalization and the production of popular music*. New Formations, 8, 99-110. 23 February, 2022, De Lawrence Wishart Database.
- The Editors of Encyclopedia Britannica. (2009). *Popular Music.* March 23, 2022. Britannica Encyclopedia. Url: https://www.britannica.com/art/popular-music
- Von Appen, R. Doehring, A. Helms, D. & Moore, A. (2015). *Song Interpretation in 21st Century Pop Music*. England: Ashgate.
- Zagorski-Thomas, S. (2014). The Musicology of Record Production. London: Cambridge University Press.

Appendix 1: Examples of graphical descriptions of music tracks as found in existing literature

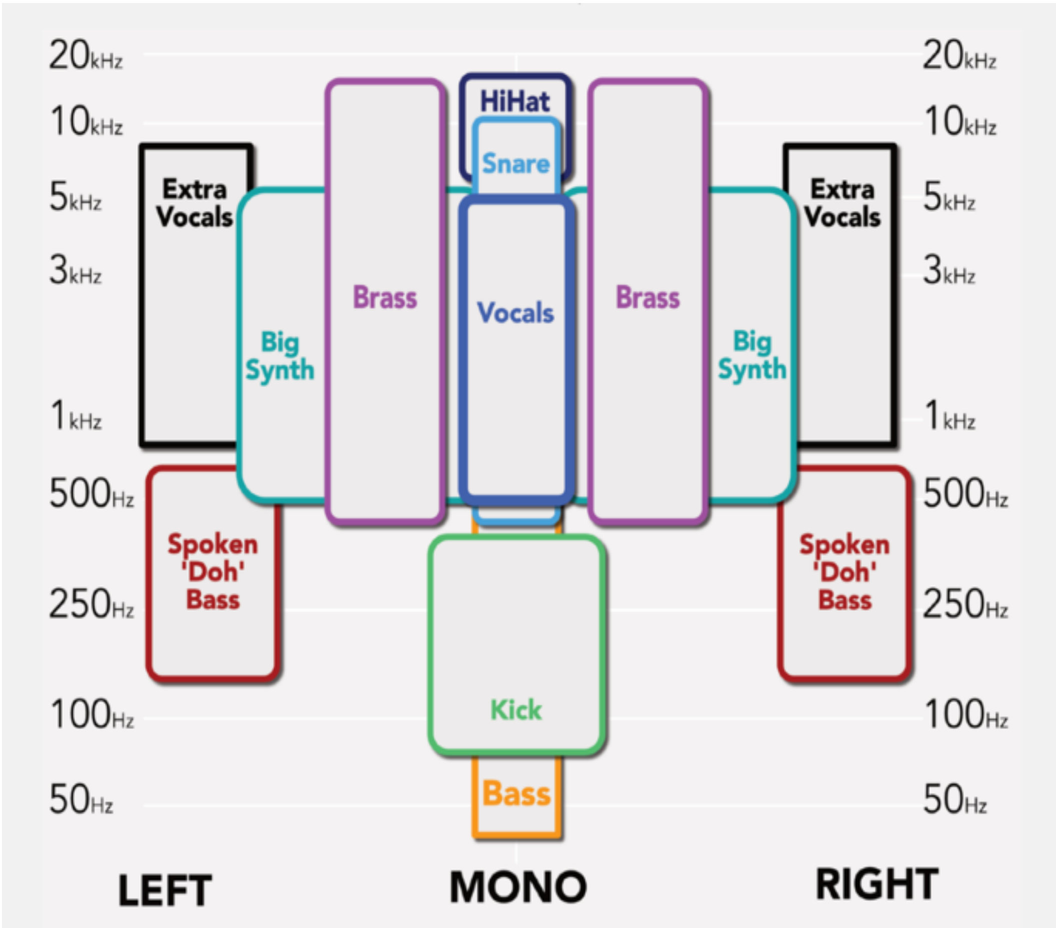

Figure 1: Extracted from "How Pros Make Hits" Frampton, T (2019). 10.

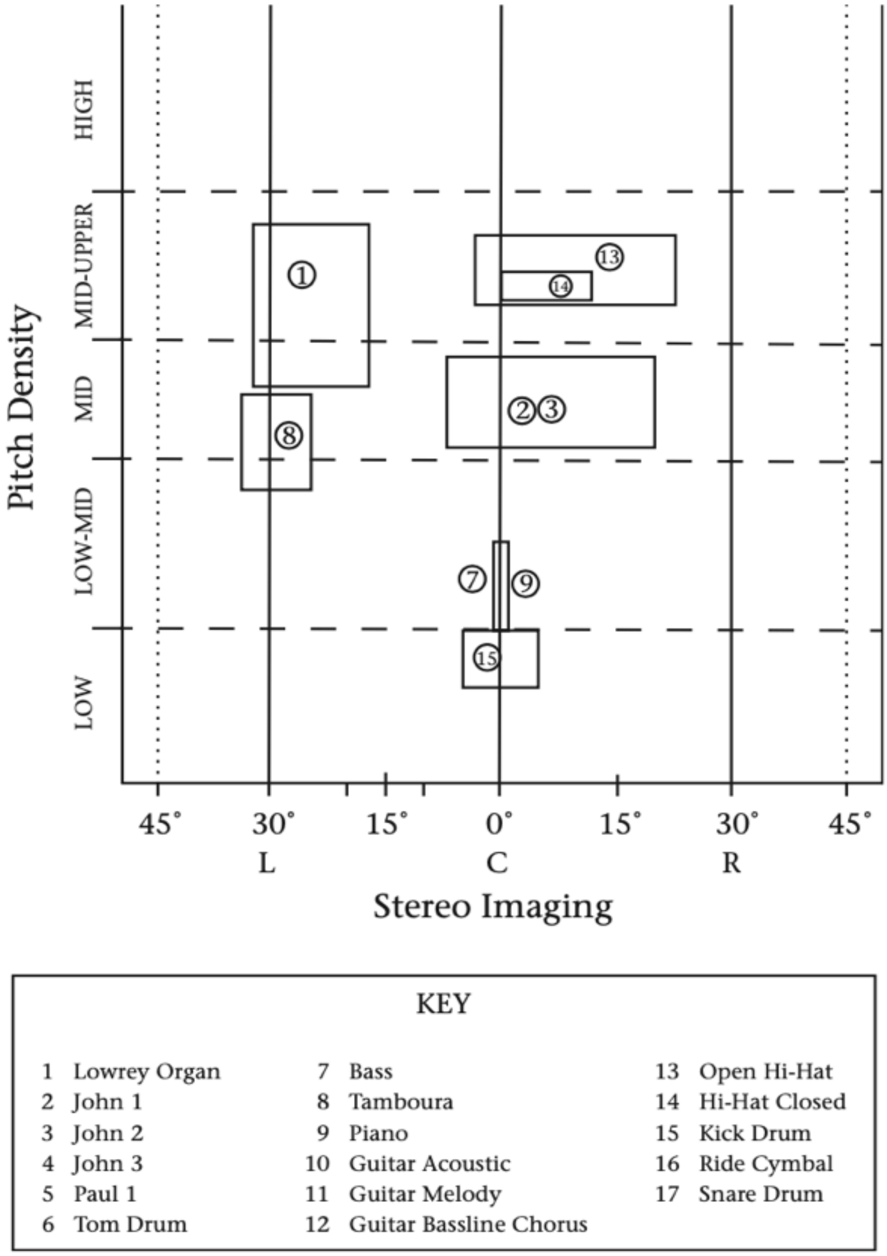

Figure 2: Extracted from "Recording Analysis: How the record shapes the song" Moylan, W (2020). 380.

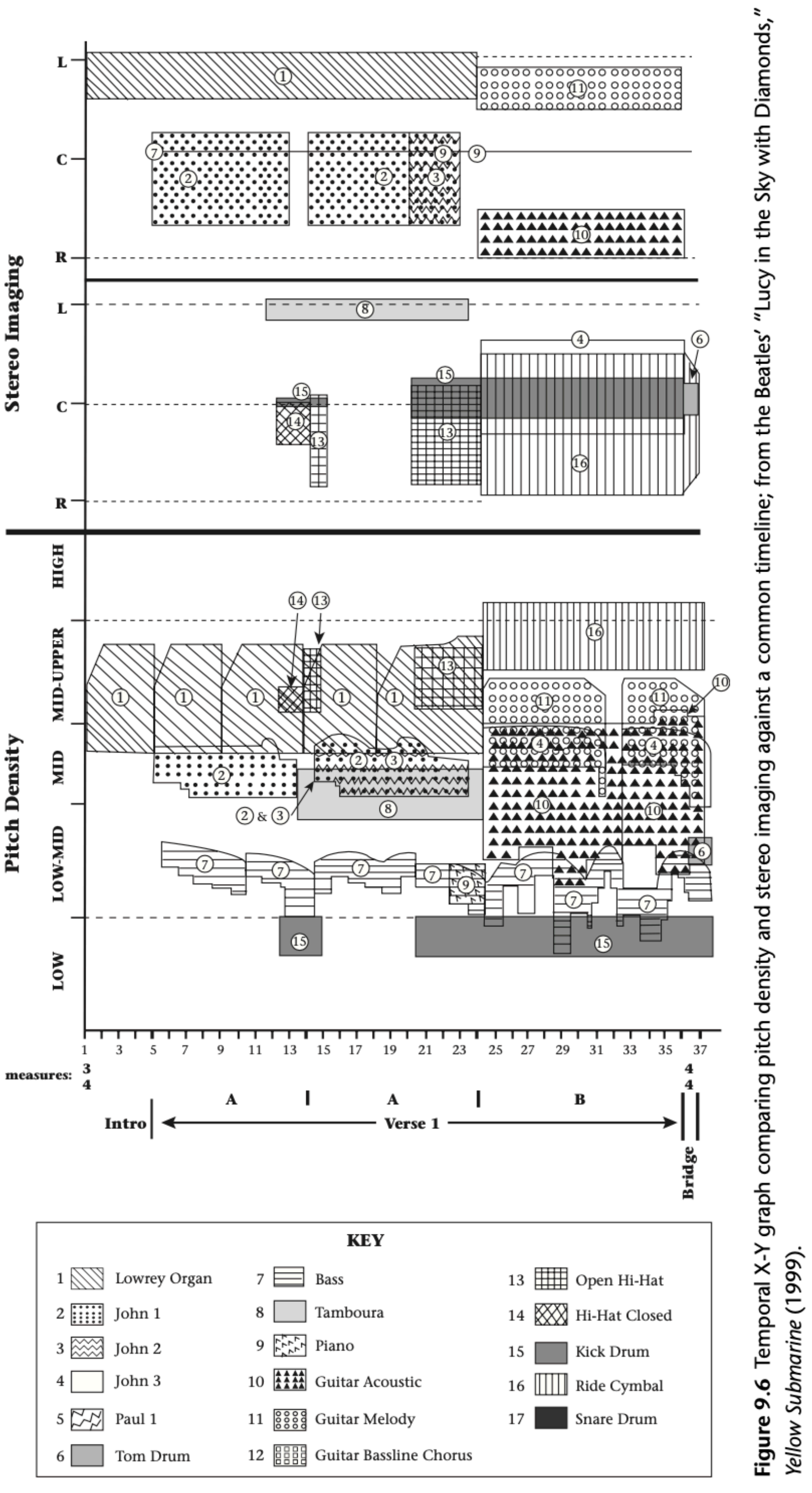

Figure 3: Extracted from "Recording Analysis: How the record shapes the song" Moylan, W (2020). 378

Appendix 2: Digital visualization tools links

- Izotope
  - RX: https://www.izotope.com/en/products/rx.html
  - Insight: https://www.izotope.com/en/products/insight.html
- Sonic visualizer: https://www.sonicvisualiser.org/
- Brainworx bx: https://www.brainworx.audio/search.html?keywords=bx_control
- Voxengo SPAN: https://www.voxengo.com/product/span/
- Fab Filter Pro Q: https://www.fabfilter.com/products/pro-q-3-equalizer-plug-in
- Youlean Loudness Meter: https://youlean.co/youlean-loudness-meter/
- Waves WLM: https://www.waves.com/plugins/wlm-loudness-meter#how-to-set-loudness-levels-for-streaming-wlm
- LEVELS: https://www.masteringthemix.com/products/levels
- Plane: https://www.pluginboutique.com/products/7190-PEEL