SCT Proceedings in Interdisciplinary Insights and Innovations. 2024; 2:276
doi: 10.56294/piii2024276

Categoría: Congreso de la Fundación Salud, Ciencia y Tecnología 2024
Original

Spectral or Dynamic Constellations as Sound Sources in Maps

Constelaciones espectrales o dinámicas como fuentes sonoras en mapas

Leandro Enrique Rodríguez[1] .

[1] National University of Quilmes (UNQ), School of Arts (EUDA), CONICET (National Scientific and Research Council), Buenos Aires, Argentina.

Citar como: Villarreal M. Cooperation for the approach of popular libraries and Wikimedia in the management of cultural projects. SCT Proceedings in Interdisciplinary Insights and Innovations. 2024; 2:280. DOI: https://doi.org/10.56294/piii2024280

Recibido: 27-04-2024 **Revisado:** 29-04-2024 **Aceptado:** 31-04-2024 **Publicado:** 03-05-2024

Editor: Rafael Romero-Carazas

ABSTRACT
Research on musical aesthetic patterns on different decades of the second part of the last twentieth century allowed specific studies based on frequency and amplitude data. This approach is carried out mainly to locate musical instruments and their relations inside the stereo image (the traditional audio virtual space), so a bidimensional map development was determined as a key part for finding spatial distribution patterns on different periods of stereo music masters -to be able to characterize timbral aesthetics from a measured perspective-. This article presents the concepts, methodology and the analysis performed, a well all their results. Hence, a new approach of musical sound sources conceptualization inside the stereo image is presented, where they are either spectral or dynamic component constellations. Additionally, this concept and approach could apply to any kind of sound sources inside a recorded signal.

Keywords: aesthetics, maps, music, sources, space, stereo.

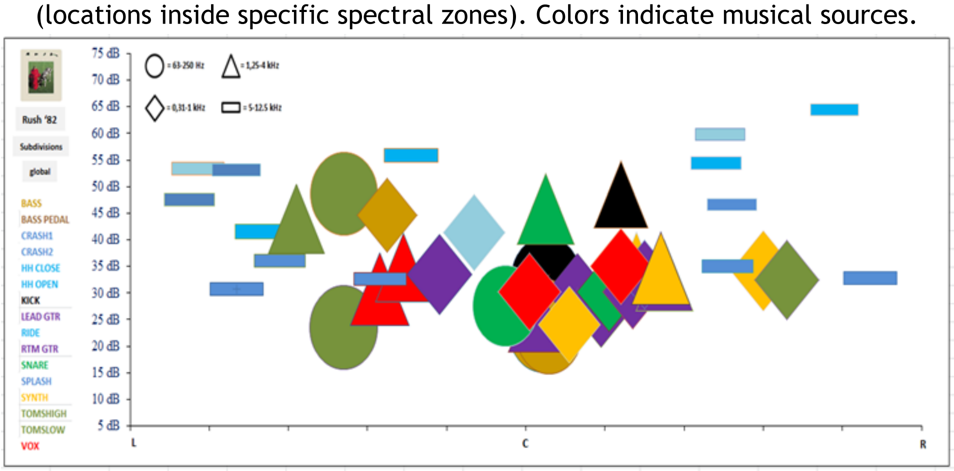
1-INTRODUCTION

Motivated by the lack of objective musical sources location methods inside the virtual space, a map construction model is presented to provide comparison possibilities for sound mixes, given the possibility to find amplitude-frequency aesthetic patterns. Hence, this article proposes a model for map constructions for sound musical sources, with the visualization goal of identifying their components inside the stereo image. This model is tested in an analysis corpus presented on this paper, based in 16 examples from the 1960's to the 1990's of the mainstream pop/rock music industry, and it's based on concepts originally developed by John Chowning [1], Albert Bregman [2], Francis Rumsey [3] [4] and Hyunkook Lee [4] [5].

This research is currently being applied on stereo masters of popular music from the recording industry golden decades (60's, 70's, 80's and 90s') in the XX century. The data were acquired through specific signal analysis resources in the frequency domain focused on amplitude -with information obtained by FFT and RMS using dBFS in L-R [6]-, and in the time domain gathering statistical spectral data by audio descriptors -with both STFT and ERB methods [7]-. The obtained values allow the construction of specific space musical source analysis tables inside the stereo image, and such information was used for building maps in which the source components can be visually identified. These values are obtained from samples of the musical instruments in the mix. Trough deep learning algorithms, specifically from Audio Source Separation [8], it is possible to obtain isolated stereo submasters from a stereo master itself, and via these more isolated tracks get specific samples for a more precise spatial analysis of several musical sources. These samples are obtained using criteria that comes both from personal professional experience and from the cultural heritage of recording, mixing and mastering techniques of popular music, concerning mainly the knowledge of common uses of the more prominent instrument spectral and spatial zones on this kind of sonic art. In other words, this is music from the mainstream industry.

The sources are analyzed full length, in a time-static averaged FFT applied to both L-R, allowing the components detection by using the higher RMS dBFS values on each specific source, as well as the use of the audio descriptors Spectral Flatness, Spectral Spread and Spectral Centroid on a time-dynamic base, allowing component classification from their temporary evolution in a statistical way [7]. All data is located on spread sheets in a third-octave frequency resolution, sending the mentioned higher RMS values to a masking competition table. This table makes a specific L-R analysis in which the higher RMS values of the first two source components of each third-octave frequency value are selected, sending this information afterwards (source name, frequency position and L-R RMS values) to a general main table, used for graph purposes (Table V). This method allows the construction of bidimensional sound maps to visually locate musical sources inside the stereo image. Two different visualization models are proposed using these resources, both using the Y axis to indicate amplitude or frequency and the X axis to defining L-R position. For each one of the components from a specific source the same color is used, for better identification for each component, as well as geometric figures with different sizes. In this way, the geometrical and size information offers what the current Y axis is not offering and vice versa, such as Figs. 1 and 2 indicate for a same musical example.

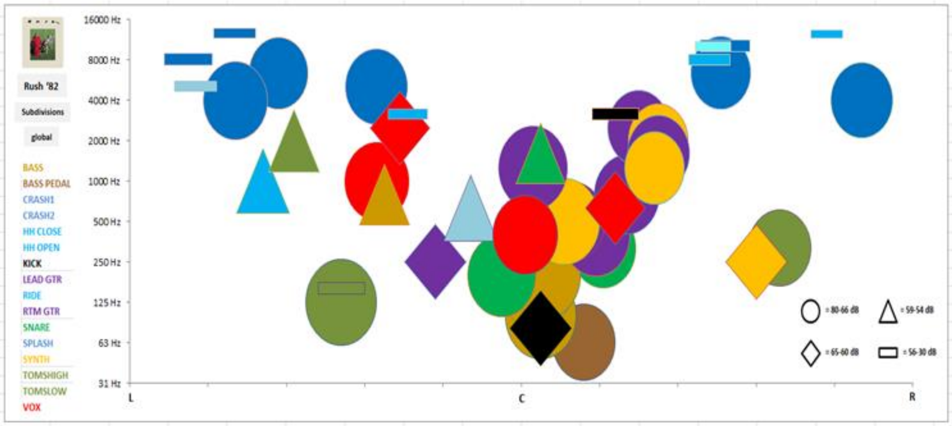
Figure 1: Map with amplitude on the Y axis, using geometric forms to indicate frequency values (locations inside specific spectral zones). Colors indicate musical sources.



Source: own elaboration

https://doi.org/10.56294/piii2023276

Figure 2: Map with frequency on the Y axis, using geometric forms to indicate amplitude values (locations inside specific dynamic range zones). Colors indicate musical sources.



Source: own elaboration

Figs. 1 and 2 shows that all source components are segregated, hence, they are spatially discrete –dispersed, located in different places inside the virtual space-. So, these components are conforming musical sources as constellations. In the next subsections it will be explained: A) how components are located, B) how a masking competition table is built in order to determine which components are going to be graphed, C) how a general table is composed with the obtained data in order to proceed with the map constructions, and D) which criteria is used for selecting, editing and using the source samples obtained by Audio Source Separation.

A- Single source component localizations

Taking Figs. 1 and 2 as an example when the lead vocal is a source composed of red spatial components, Table I shows their frequency and azimuth positions.

TABLE I: EXAMPLE OF FREQUENCY AND AZIMUTH POSITIONS

Freq	**Freq Pos**	RMS L	RMS R	L-R Diff	**L-R Pos**
220.7 Hz	**200 Hz**	-29.04 dBFS	-29.88 dBFS	0.84	**-8.38**
436.0 Hz	**400 Hz**	-30.31 dBFS	-30.43 dBFS	0.11	**-1.14**
646.0 Hz	**630 Hz**	-36.34 dBFS	-33.87 dBFS	-2.46	**24.62**
952.8 Hz	**1000 Hz**	-28.89 dBFS	-32.60 dBFS	3.71	**-37.12**
2476.3 Hz	**2500 Hz**	-33.28 dBFS	-36.41 dBFS	3.14	**-31.36**
3111.5 Hz	**3150 Hz**	-47.11 dBFS	-44.20 dBFS	-2.91	**29.13**
5065.6 Hz	**5000 Hz**	-50.60 dBFS	-48.84 dBFS	-1.76	**17.57**

On Table I, the Freq column indicates absolute Hz value obtained by averaged FFT, while Freq Pos -used for map frequency position of each component- uses a round out to a third octave resolution value, due to achieve a simple method closer to the critical bands [9]. Also, RMS L & R indicate dBFS values for each channel, and L-R Diff is the subtraction between L and R. This method is based on the ITD e ILD concepts (Interaural Time Difference and Interaural Intensity Difference, respectively) [6], as well the as the precedence effect and the frequency radiation patterns of acoustic sources. 10 dB differences between channels are both used on ILD for coincident stereo microphone takes [5] and the precedence effect [9] to settle a source full L or full R in the azimuth position, as well as broadband criteria concerning musical sources contemplating loudness curves [9]. In this way, the 10 dB value obtained as L-R is assigned as full R, and -10 dB as full R. This value is taken by L-R Pos -used of the

https://doi.org/10.56294/piii2024276

map azimuth position each component-, multiplying it by -10 to adjust a scale in which 0 is the center (C), -100 full L and 100 is full R. Using this principle to determine azimuth X position, the frequency-based maps use Hz and the amplitude-based ones use dB to define Y position on both map models.

TABLE II: EXAMPLE OF FREQUENCY AND AMPLITUDE POSITIONS

Freq	**Freq Pos**	**RMS L**	**RMS R**	**RMS Avg**	**RMS Pos**
220.7 Hz	**200 Hz**	**-29.04 dBFS**	**-29.88 dBFS**	**-29.46**	**29.46**
436.0 Hz	**400 Hz**	**-30.31 dBFS**	**-30.43 dBFS**	**-30.37**	**30.37**
646.0 Hz	**630 Hz**	**-36.34 dBFS**	**-33.87 dBFS**	**-35.10**	**35.10**
952.8 Hz	**1000 Hz**	**-28.89 dBFS**	**-32.60 dBFS**	**-30.74**	**30.74**
2476.3 Hz	**2500 Hz**	**-33.28 dBFS**	**-36.41 dBFS**	**-34.84**	**34.84**
3111.5 Hz	**3150 Hz**	**-47.11 dBFS**	**-44.20 dBFS**	**-45.46**	**45.46**
5065.6 Hz	**5000 Hz**	**-50.60 dBFS**	**-48.84 dBFS**	**-49.72**	**49.72**

SOURCE: OWN ELABORATION

On Table II, the RMS Avg column indicates the average amplitude values between L and R, while RMS Pos -used for the map amplitude position each component- obtains the positive value from RMS Avg multiplying it by -1, to be able to settle positive values on a Y axis, as Fig. 1 shows. Also, a geometrical code based on specific spectral zones is used for frequency values for this kind of maps (Fig. 3), as well as specific amplitude ranges applies other geometrical code for maps with frequency on the Y axis (Fig. 4).

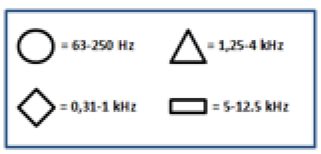


Fig. 3: Spectral range zones: geometrical code for frequency value notion in maps with amplitude on the Y axis. This code is used for spectral constellations. Source: own elaboration.

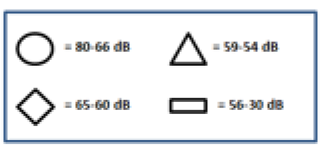
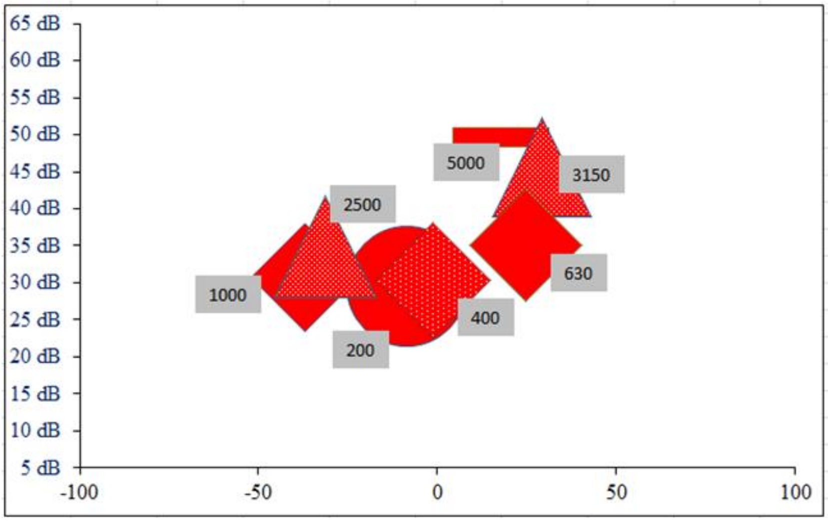


Fig. 4: Dynamic range zones: geometrical code for amplitude value notion in maps with frequency on the Y axis. This code is used for dynamic constellations. Source: own elaboration.

These geometrical codes -besides giving more visual information concerning the nature of the components- are very useful in showing their spatial distribution with more detail, as well as the interaction with components of other sources. Using the data from Tables I and II with the geometrical codes of Figs. 3 and 4, a constellation of components is obtained for this source: it is located in both map models, the one with amplitude on the Y axis (Fig. 5) and the one which uses frequency (Fig. 6), using the X axis to determine azimuth position in both models, with range -100 (full L) to 100 (full R), using 0 as the Center (C).

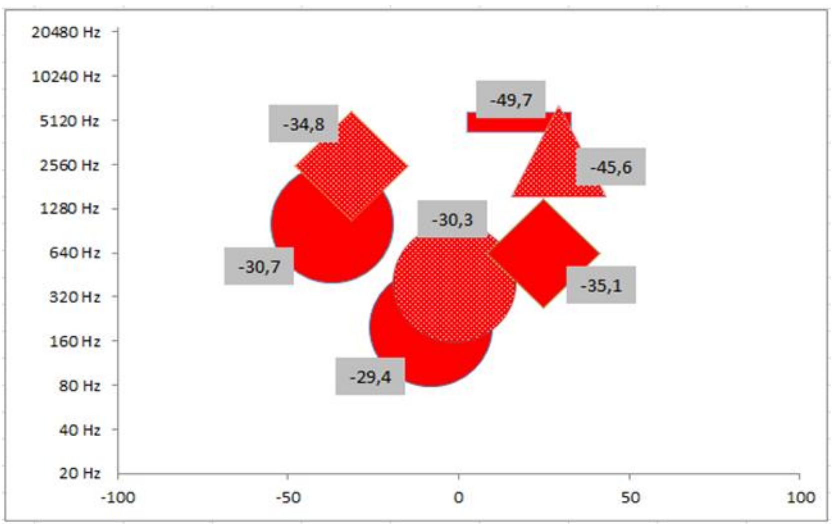
Figure 5: Example of a source constellation of 7 components, using the model with amplitude on the Y axis and a geometrical code for frequency ranges. Azimuth is defined as -100 = full L, 0 = C and 100 = full R. This is defined as a spectral source constellation because its components define frequency positions for the entire sound source.

https://doi.org/10.56294/piii2023276



Source: own elaboration

Figure 6: Example of a source constellation of 7 components, using the model with frequency on the Y axis and a geometrical code for amplitude ranges. Azimuth is defined as -100 = full L, 0 = C and 100 = full R. This is defined as a dynamic source constellation because its components define amplitude positions for the entire sound source.



Source: own elaboration

The FFT obtained values are classified as valid if the analysed source has the desired isolation level respecting other sources. This is relevant because defines the fact that the inharmonic content of its timbre has no information of other instruments, so running spectral analyses afterwards is pertinent.

This task is assigned to Spectral Flatness [10], an audio descriptor that brings the capability of classifying audio signals as tonal or percussive depending on a statistical evolution of its spectrum, in which coefficient 0 indicates the spectrum of a perfect periodical signal and 1 defines pure flat spectrum. A specific criterion is used, =>0.9 for percussive sounds (for example kick, snare, toms, cymbals) and <0.9 for harmonic ones (vocals, bass, guitars, keyboards). For the example already mentioned on Tables I and II, Spectral Flatness coefficient indicates 0.89.

https://doi.org/10.56294/piii2024276

Spectral Spread is another audio descriptor used in this model -is also known as Spectral Standard-Deviation [10]-. It represents the spectral movement around the mean value of the full frequency response analysed on a time function. This definition allows thinking in a link of the ADSR concept to a statistical spectral evolution, specifically for detection of frequency values on the signal attack (in the first 10 frames, 0-5 ms range) and the signal sustain afterwards, using a 180 frames per second resolution. As another statistical complement in the time function, Spectral Centroid (which represent the evolution of the gravity centre of the spectrum) [10] is another utilized audio descriptor, specifically as a method for obtaining minimum, maximum and average temporally evolution to highlight a spectral zone not detected by Spectral Spread. For the example showed on Tables I and II, the average Spectral Centroid is defined on 3150 Hz (rounded out at third octave resolution from 3100 Hz) between 2600 Hz and 3600 Hz as minimum and maximum values. This allows the construction of a new table (Table III), which indicates at least 2 audio descriptors validating 2 source components via statistical spectral evolution, not average static as the use of standard FFT in this methodology.

TABLE III: AUDIO DESCRIPTORS ATTACHED TO COMPONENTS

Freq	RMS L	RMS R	Audio Descriptor
200 Hz	-29.04 dBFS	-29.88 dBFS	-
400 Hz	-30.31 dBFS	-30.43 dBFS	-
630 Hz	-36.34 dBFS	-33.87 dBFS	-
1000 Hz	-28.89 dBFS	-32.60 dBFS	**Spectral Spread**
2500 Hz	-33.28 dBFS	-36.41 dBFS	-
3150 Hz	-47.11 dBFS	-44.20 dBFS	**Spectral Centroid**
5000 Hz	-50.60 dBFS	-48.84 dBFS	-

SOURCE: OWN ELABORATION

B- Masking competition table

Values from Table III are obtained from the highest amplitude detected by the FFT analysis. As it can be seen in Figs. 1 and 2, not all the components of this source are located with the rest of the constellations of other sources building a full ensemble on the same musical example. Of the 7 detected components of the source described in Table III (in 200 Hz, 400 Hz, 630 Hz, 1000 Hz, 2500 Hz, 3150 Hz and 5000 Hz) only 5 passed to the maps (components in 3150 Hz and 5000 Hz were taken out). This is because a masking competition table selects the first 2 components with more average L-R amplitude in each third octave slot, as it can be seen on the example of Table IV.

TABLE IV: EXTRACT FROM THE MASKING COMPETITION TABLE

Freq	Source	RMS L	RMS R	RMS Avg
2500 Hz	**LEAD GTR**	-34.94 dBFS	-31.91 dBFS	**-33,42**
2500 Hz	**VOX**	-33.28 dBFS	-36.41 dBFS	**-34,84**
2500 Hz	**SYNTH**	-36.43 dBFS	-33.60 dBFS	-35,01
2500 Hz	**RTM GTR**	-36.78 dBFS	-45.21 dBFS	-40,99

source: own elaboration

Table IV shows 4 components in 2500 Hz, part of the sources detailed in the Source column, along with their L-R RMS amplitude values in dBFS and the average value between them, in the RMS Avg column.

The first 2 components with more amplitude qualify as map components in each third octave value: these 2 passes to a general table, used for map construction. Sometimes 3 places are assigned in this table when relevant components are detected around L, C and R (3 common location instances in the stereo image), always when there are no components with more amplitude at the sides of the third octave value when these components are detected. However, it is possible to insert more components in certain cases, such as in empty zones values (for the lack of source components in that frequency slots) or well for drastic timbre changes for a single source. So, a strategy is established when a value can be removed or inserted in a spectral zone when no components are located, thanks to the existence of one in a near and relevant position, in 3 possible cases:

1) For example, if there are no components in 800 Hz but in 630 Hz (adjacent value) 3 components are detected, with prominent amplitude in L, C and R- one slot in 800 Hz is deleted to insert one for the 630 Hz value on the general table.

2) If components adjacent to a slot are assigned already and still the value is empty on the general table, the component placed is the one with more intensity detected in all sources by FFT analysis. So, more than 4 source components exist in these situations, making this case more common than 1).

3) If a single source has a drastically timbre change, a component is created for each case (for example hh open / hh close -open hi hat / close hi hat-, or gtr arp / gtr rtm –arpeggio guitar / rhythm guitar, etc.). This component qualifies both to fill an empty slot on the general table and to compete freely with other source components in the masking competition table.

C-General Table

Table V shows a general source components table example, built from a masking completion table, with third octave frequency values from 63 Hz to 12.5 kHz. The components define their belonging to a source by color, name and spectral range occupied inside the spectrum. Table V also shows values of the source components in maps of Figs. 1 and 2. With the goal to get better visual legibility, two specific size criteria are used for each source component. When the Y axis is amplitude based, “the higher the frequency the smaller the component” is used (this somehow links to the wavelength inversely proportional to frequency concept [11]). On the other hand, when the Y axis is frequency based the amplitude of the component defines its size, using a positive scale. For example, if -30 dBFS is the component amplitude (assuming standard 16 bits 96 dBFS dynamic range) applies 96-30=66, so this positive value defines its size on the map.

TABLE V: GENERAL SOURCE COMPONENTS TABLE EXAMPLE

Hz Pos	SOURCE component	RMS L (dBFS)	RMS R (dBFS)	RMS Pos	L-R Pos
63	BASS PEDAL	-28.3	-26.6	27.4	16
63					
80	KICK low	-32.4	-32.9	32.6	5
80					
100	BASS low1	-24.9	-24.4	24.5	5
100					
125	TOMSLOW low	-21.2	-25.8	23.5	-46
125					
160	BASS low2	-23.3	-22.9	23.1	4
160	TOMSHIGH low	-22.5	-27.2	48.8	-46
200	BASS low3	-22.7	-22.0	22.3	6

https://doi.org/10.56294/piii2024276

200	SNARE low	-27.2	-27.8	27.4	-5
250	RTM GTR low	-32.6	-34.8	33.6	-22
250	SYNTH low	-37.2	-31.2	34.1	60
315	SNARE mid	-31.3	-29.2	30.2	21
315	TOMSHIGH mid	-35.7	-29.1	32.4	66
400	LEAD GTR mid1	-28.1	-26.2	27.1	19
400	VOX mid1	-30.3	-30.4	30.3	1
500	RTM GTR mid	-31.0	-29.6	30.2	13
500	SYNTH mid	-24.7	-23.6	24.1	11
630	RIDE mid	-40.7	-42.1	41.4	-13
630	VOX mid2	-36.3	-33.9	35.1	24
800	BASS mid	-42.9	-46.5	44.7	-35
800	LEAD GTR mid2	-31.8	-29.0	30.3	27
1000	VOX midhigh1	-28.9	-32.6	30.7	-37
1000	HH CLOSE midh1	-38.3	-44.9	41.6	-66
1250	RTM GTR midh1	-25.7	-26.0	25.8	3
1250	SYNTH midhigh1	-36.5	-33.0	34.7	34
1600	SNARE midhigh1	-46.1	-45.6	45.8	5
1600	LEAD GTR midh1	-35.1	-31.5	33.3	35
2000	SYNTH midhigh2	-36.4	-33.6	35.0	28
2000	TOMSHIGH midh	-41.0	-46.9	43.9	-58
2500	VOX midhigh2	-33.3	-36.4	34.8	-31
2500	LEAD GTR midh2	-34.9	-31.9	33.4	30
3150	KICK midhigh	-48.8	-48.5	48.6	24
3150	HH close midh2	-55.4	-58.4	55.9	-29
4000	CRASH1 high1	-27.1	-34.4	30.7	-73
4000	CRASH2 high1	-37.3	-28.5	32.8	87
5000	CRASH1 high2	-30.8	-34.6	32.7	-37
5000	RIDE high1	-48.3	-58.6	53.4	-83
6300	CRASH1 high3	-33.2	-39.2	36.1	-62
6300	CRASH2 high2	-37.7	-32.6	35.1	51
8000	HH OPEN high1	-56.8	-52.0	54.4	48
8000	SPLASH high1	-41.9	-53.2	47.5	-85
10000	CRASH2 high3	-49.1	-44.0	46.5	52
10000	RIDE high2	-62.3	-57.3	59.8	49
12500	HH OPEN high2	-68.4	-60.5	64.4	78
12500	SPLASH high2	-48.4	-57.7	53.1	-73

SOURCE: OWN ELABORATION

D-Selecting and editing audio samples from files

Clearly there are many ways to do this, and several of them were tested. The most effective method found is based on what to represent a start point, and two different paths emerged from this question as a point of view. One option is by thinking the sources globally, using an average implementation of all the song/piece for each relevant source (for example on a kick drum, averaging spectrums of all the kicks in the song or taking only one as the most representative of the sound of the kick). Another method is by thinking the sources structurally, using the internal parts of the song/piece (the kick from the verse, the kick from the chorus, the kick from the bridge, etc.). This start point is important because it sets a scenario based on applying the same methodology on all the sources to analyze, to represent a “global map of the song” or well a structural, “specific song part map”.

For the examples on the maps developed in the analysis corpus in section II, both global and structural criteria were used, with each case specified in the figure descriptions. Once this decision is taken, it’s important to define the suitable length of the samples to extract after applying the Audio Source Separation process. Here 4 factors were defined:

1) Interference between sources: Samples with minimum sustain and/or fast release (such as finger snaps, for example) cannot have a long duration because other sources mask the end of the sample, though a very low audible contamination from other instruments. In these cases, the sample length oscillates

between 50 ms and 300 ms because this time range somehow frames all the ADSR envelopes of the sources pretended to analyse.
2) Amplitude data: the sample needs to have 50 ms as minimum time length for obtaining relevant RMS values from L-R dBFS in most FFT cases.
3) Audio descriptors evaluation: if a source has substantial information below 300 Hz, the sample needs to be => 500 ms to allow descriptors such as Spectral Spread to give data in the low frequency zone, something that not occurs with sources of less durations.
4) The length of a musical note depends on tempo, the chronometric density and the timbral function of the source on the textural discourse [12]. Logically, this concept offers cases when the sample crosses the 1 s time threshold, but an arbitrary length of 999 ms is established as maximum, considering the source in these cases from its attack (the first 5 ms / 10 ms since the beginning of the envelope on the signal).
With these considerations, for the entire source analysis corpus (from 88 different sources) an average sample length of 477 ms is obtained. The percussive sources use to be the shorter ones (50 ms to 200 ms) while basses, guitars, keyboards, and vocals are longer (200 ms to 999 ms). Additionally, all sources are treated with the shorter possible fades –ins and outs-, from 0.3 ms to 1 ms to achieve beginnings and endings with no additional analysis artefacts. As another part of this methodology, all the FFT analysis is made using Blackman-Harris windows of 8192 sample size, and the audio descriptors setup (both on STFT and ERB methods) is defined on Hanning windows with 4096 sample size. These different window type and size configurations for the spectral data are due to the search of more precision on the obtained static and averaged values detected by standard FFT [13], and on the other hand, more precision on the temporary evolution on the dynamic values obtained by the audio descriptors, both on STFT and ERB methods. This situation established a time-frequency compromise relationship assigned to each obtained source component.

2-IMPLEMENTATION

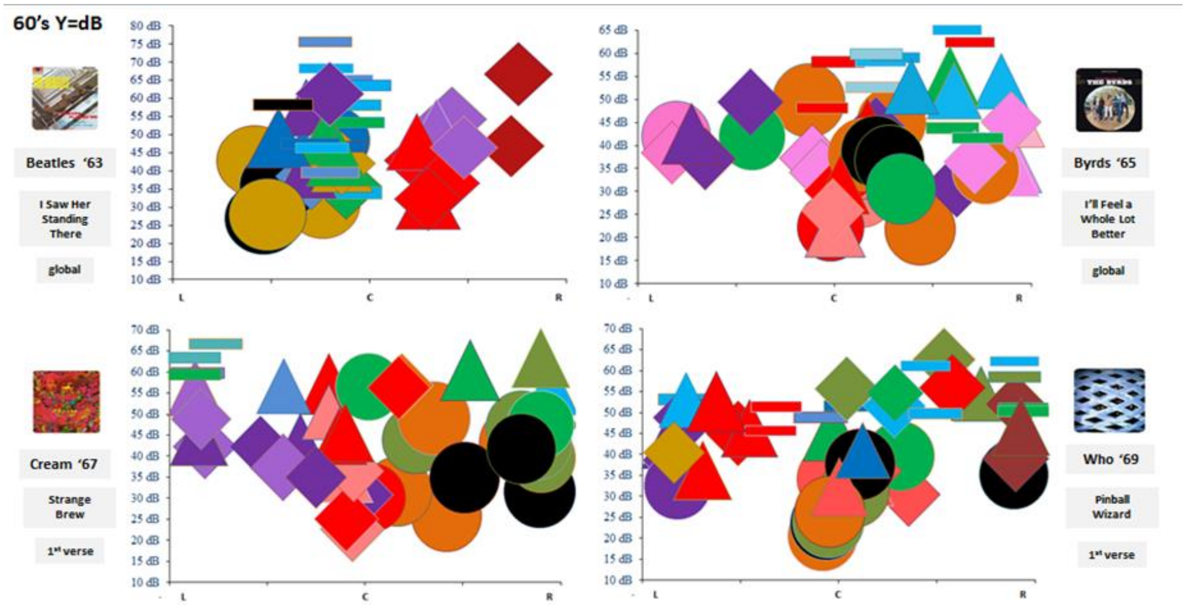
This section shows all maps made with the described methodology. It shows all obtained maps in a decade partner comparison for a quick sight check, resulting in an inevitable visual comparison among them.
Each example of this decade-based corpus also can be global or structural in each case, as described below.

1960’s:
-1963: The Beatles, I Saw Her Standing There, global.
-1965: The Byrds, I’ll Feel a Whole Lot Better, global.
-1967: Cream, Strange Brew, structural (1st verse).
-1969: The Who, Pinball Wizard, structural (1st verse).
1970’s:
-1972: Yes, Close to the Edge, structural (1st verse).
-1975: Electric Light Orchestra (ELO), Evil Woman, structural (1st verse).
-1977: Pink Floyd, Dogs, global.
-1979: The Police, Message in a Bottle, global.
1980’s:
-1982: Rush, Subdivisions, global.
-1984: U2, Pride (In The Name Of Love), structural (1st chorus).
-1986: Queen, A Kind of Magic, global.

https://doi.org/10.56294/piii2024276

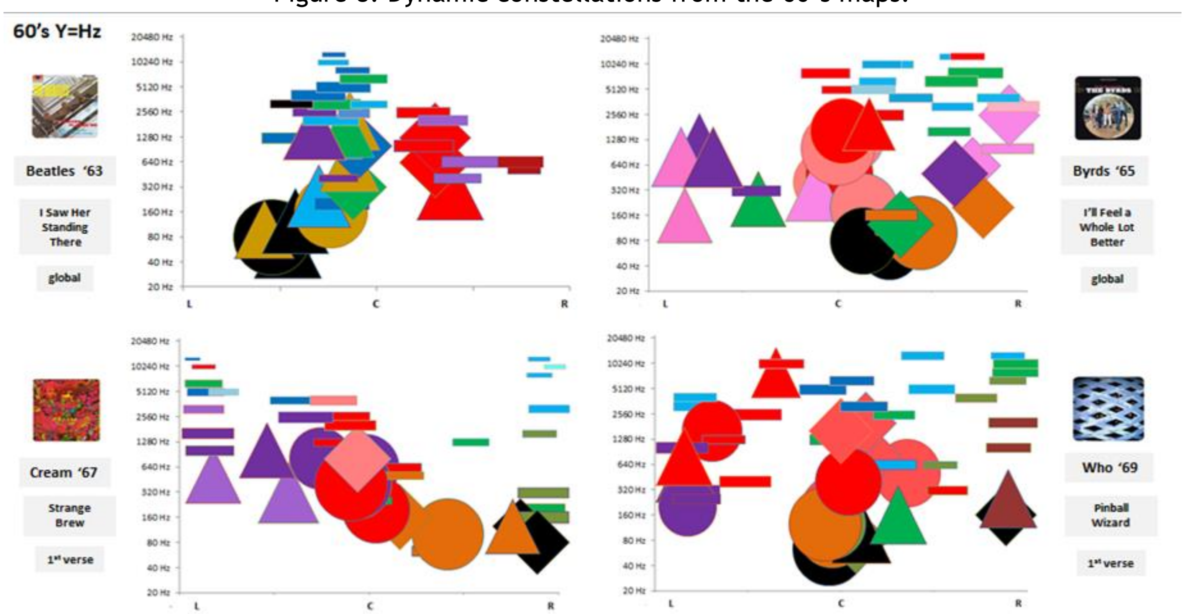
-1989: The Cure, Pictures of You, structural (1st verse)
1990's:
-1991: Nirvana, Smells Like Teen Spirit, structural (1st chorus)
-1993: Depeche Mode, I Feel You, global.
-1995: Radiohead, The Bends, global.
-1998: The Smashing Pumpkins, Ava Adore, structural (1st chorus and solo).

Figure 7: Spectral constellations from the 60's maps.



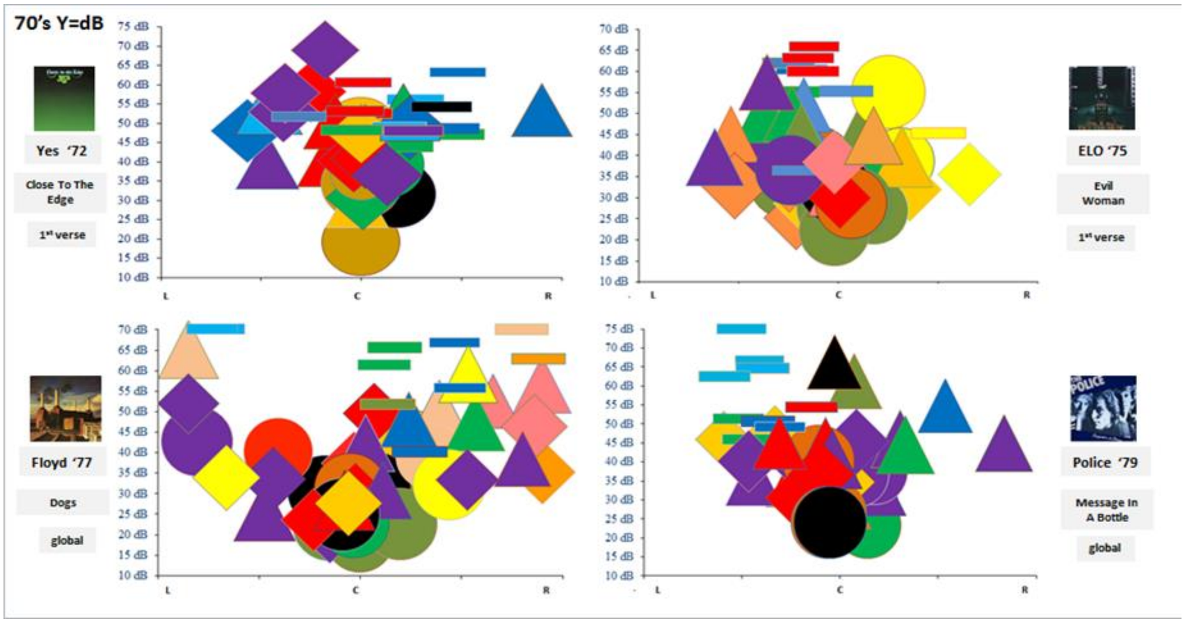
Source: own elaboration

Figure 8: Dynamic constellations from the 60's maps.



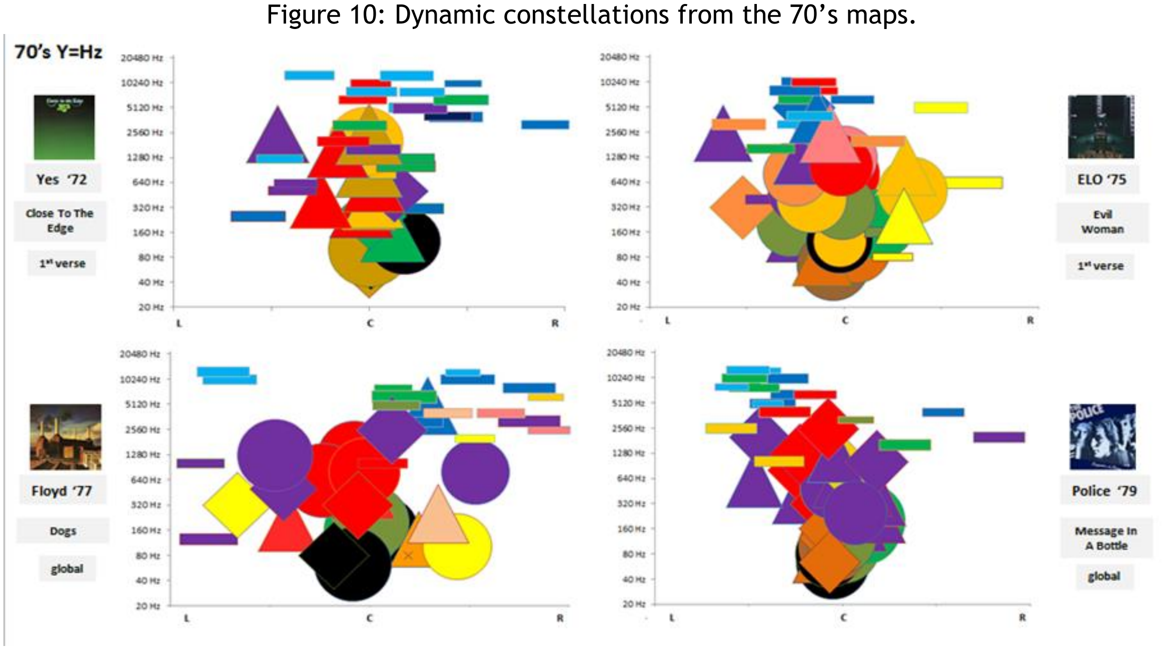
Source: own elaboration

Figure 9: Spectral constellations from the 70's maps.



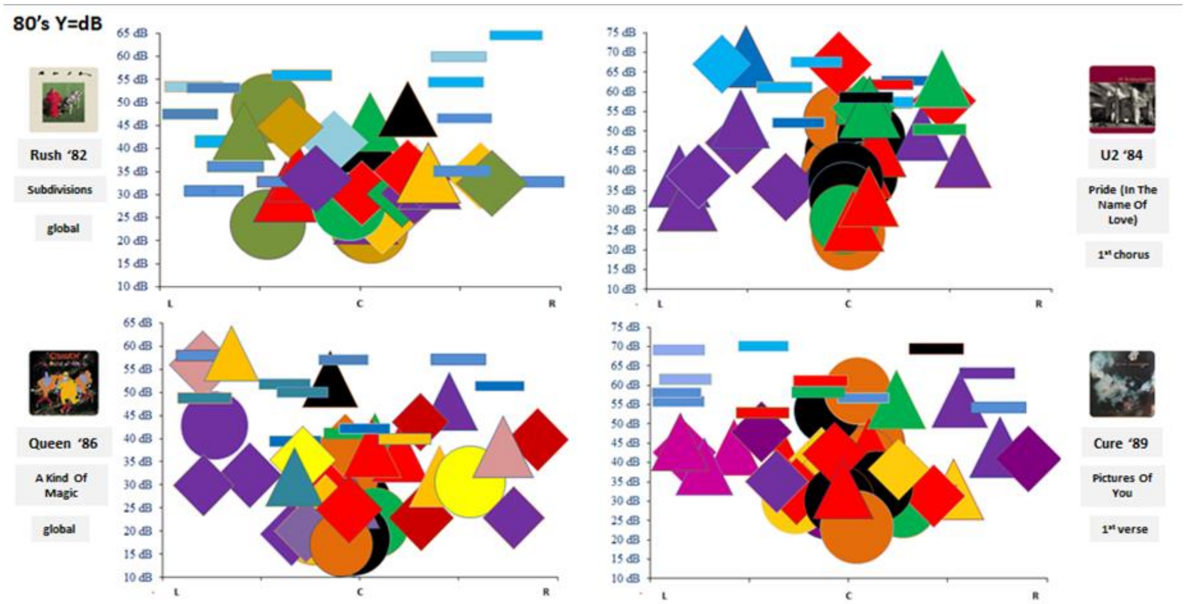
Source: own elaboration

Figure 10: Dynamic constellations from the 70's maps.



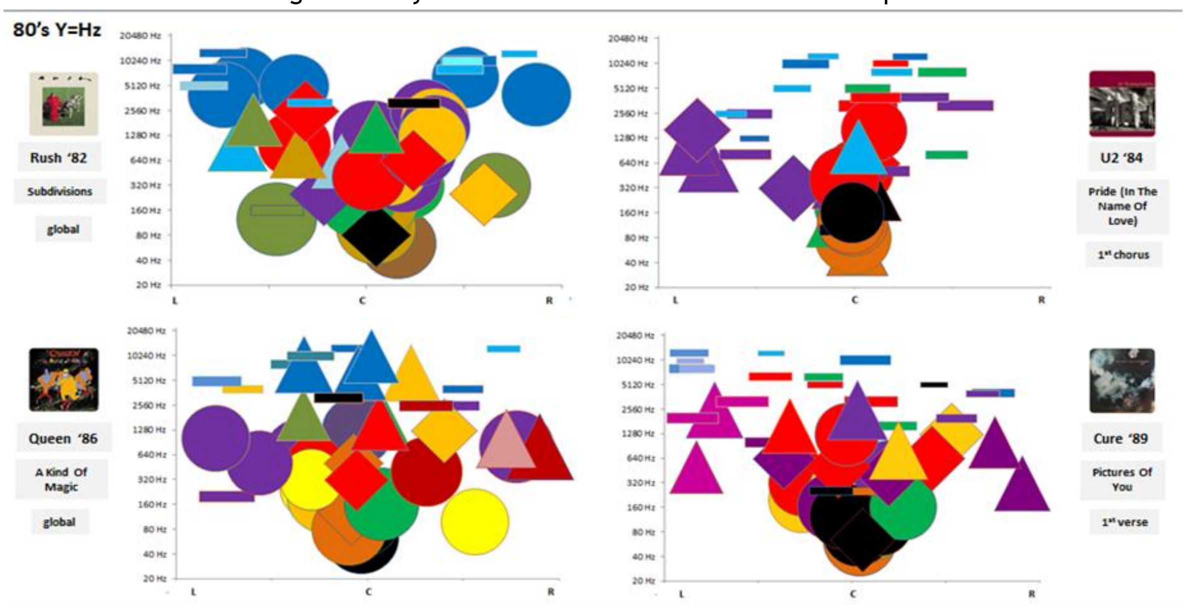
Source: own elaboration

Figure 11: Spectral constellations from the 80's maps.



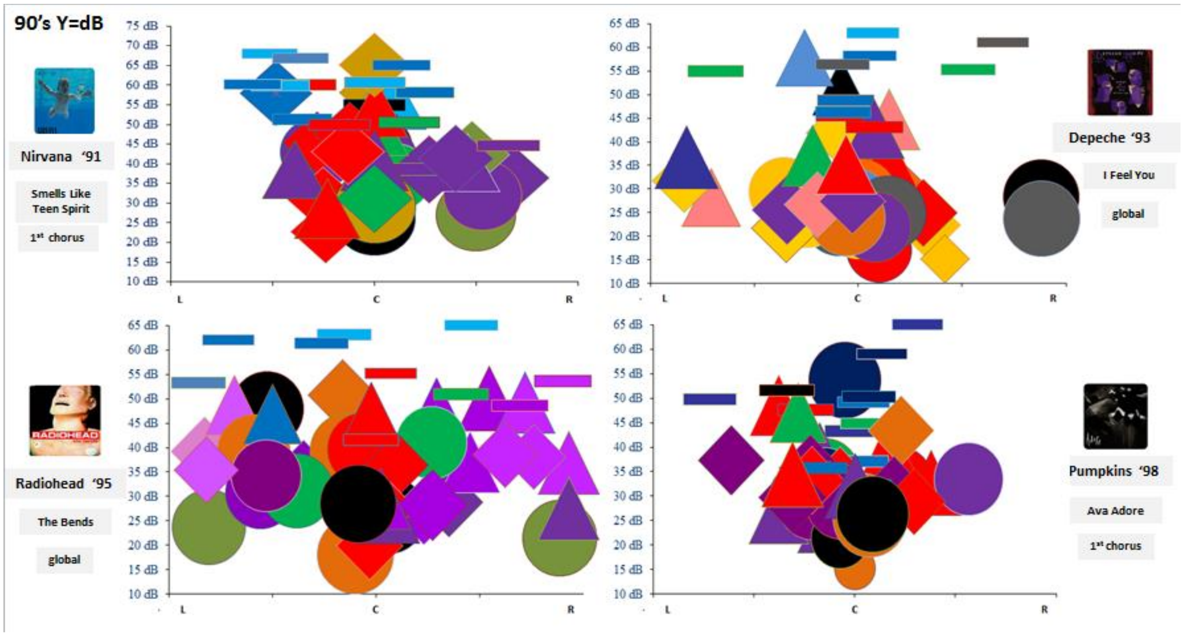
Source: own elaboration

Figure 12: Dynamic constellations from the 80's maps.



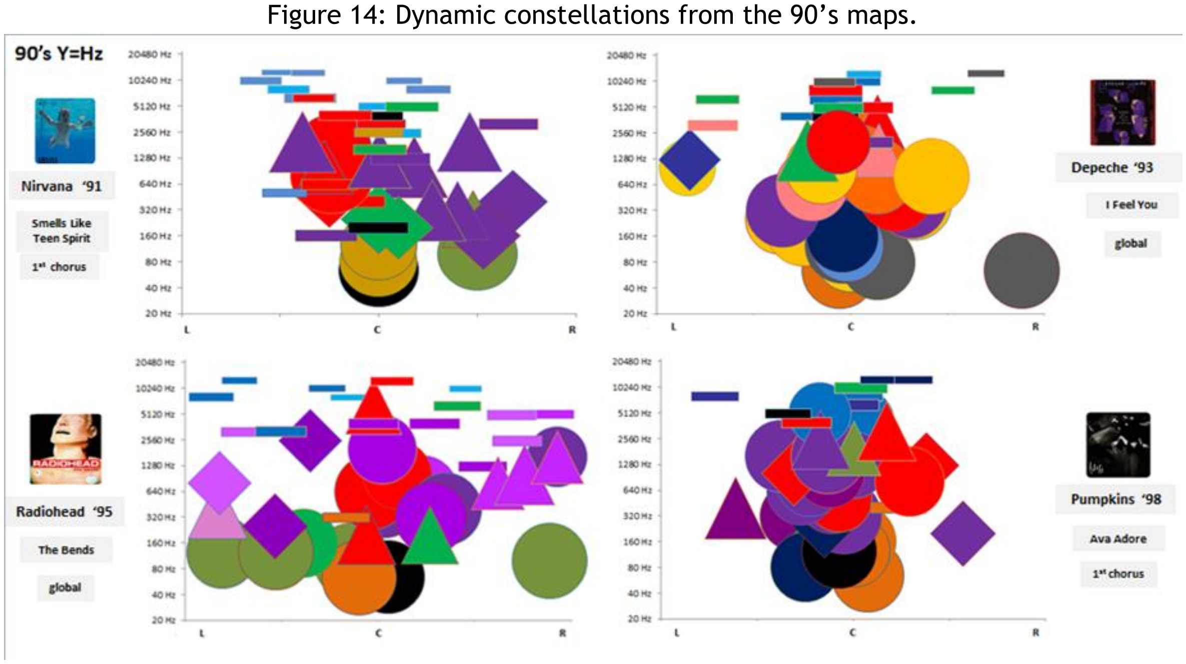
Source: own elaboration

Figure 13: Spectral constellations from the 90's maps.



Source: own elaboration

Figure 14: Dynamic constellations from the 90's maps.



Source: own elaboration

Certain observations are gathered in the next assertions, listed in order of simplicity (from the simplest to the most complex):

- Mixes are not perfectly symmetrical around the centre in reference to the spatial source components distribution, and spectral charge between what is located on the left and what is located on the right has different information in quantity, type, and location of the source components.

-Most of low frequency components tend to be near the centre, most of the high frequency ones near the sides, and most of spectral midrange components are located between low ones and high ones, in full agreement with previous studies made on massive musical discourses [14].
-The idea of low frequency components as "big ones" and higher frequency components as "smaller ones" in maps with amplitude on the Y axis in relation to the way they define their size in the maps with frequency on the Y axis (by amplitude positive values) gives similar visual results. This component size is somehow equivalent in both map models.
-In the 60´s examples, mixing with sources in stereo starts from nullity of components in the centre to minimum balance but with wide L-R disposition. In the 70´s, more mono commitment and better technology shows mostly more balanced and legible components around the centre, with narrow stereo images. In the 80´s most of the images became wider again, perhaps more than ever, and the 90´s again mostly condense components around the centre. All these definitions could indicate that -in the aforementioned spatial features- the 60´s perhaps are in a way like the 80´s and the 70´s like the 90´s, establishing a possible aesthetic tendency around the distribution of constellations of source components in the last 4 decades of the XX century, in the music industry context.
-It is relevant to point out that the exceptions of this tendency are the Pink Floyd and Radiohead examples, maps from these two artists are similar. These two rock bands are commonly aesthetically linked, but in this particular song there are substantial different timbral and chronometric attributes, although the source components are spatially similar. This would also lead to the search of aesthetic patterns [15] across decades, with not specifically the same sonic palette between musical samples.

3-CONCLUDING REMARKS

In all cases the sound sources are constellations of spatial components. In some cases, in a prominent way, this means with a higher spatial separation and/or spectral & dynamic dispersion. From results obtained it is possible to draw attention to several features, as preliminary conclusions:

- Besides obtaining measurable aesthetic patterns, by using these maps it is possible to give objectivity to the virtual sound space, taking as a model the subjective scene paradigm of Rumsey and the source components conceptualization [16], and somehow start to settle some objective way to implement it.
- This model could be used for education, analysis, composition, and other applications outside music, such as discipline fields with the need to catalogue sound as a discourse composed by component constellations, this is, sound as something generated by sources with their spectral or dynamic components spatially segregated.
- The role of mixing in musical recordings is often below composing and producing material. It is believed that the mixing engineers can somehow hear this kind of sound organization [17]: this will be tested in specific experiments as the final step for this current PhD research [18].

REFERENCIAS BIBLIOGRÁFICAS

1. Chowning JM. The simulation of moving sound sources. Journal of the audio engineering society. 1971 Jan 1;19(1):2-6.
2. Bregman AS. Auditory scene analysis: Hearing in complex environments. MIT press. 1994
3. Rumsey F. Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. Journal of the Audio Engineering Society. 2002 Sep 15;50(9):651-66.

https://doi.org/10.56294/piii2023276

4. Lee H, Rumsey F. Level and time panning of phantom images for musical sources. Journal of the Audio Engineering Society. 2013 Dec 20;61(12):978-88.
5. Lee H, Johnson D, Mironovs M. An interactive and intelligent tool for microphone array design. In Audio Engineering Society Convention 143 2017 Oct 8. Audio Engineering Society.
6. Basso G. Análisis espectral: la transformada de Fourier en la música. Ediciones al Margen; 2001.
7. Peeters G, Giordano BL, Susini P, Misdariis N, McAdams S. The timbre toolbox: Extracting audio descriptors from musical signals. The Journal of the Acoustical Society of America. 2011 Nov 1;130(5):2902-16.
8. Makino S, editor. Audio source separation. Berlin, Germany: Springer; 2018 Mar 1.
9. Zwicker E, Fastl H. Psychoacoustics: Facts and models. Springer Science & Business Media; 2013 Mar 14.
10. Caetano M, Saitis C, Siedenburg K. Audio content descriptors of timbre. Timbre: Acoustics, perception, and cognition. 2019:297-333.
11. Miyara F. Acústica y sistemas de sonido. Universidad Nacional de Rosario; 2003.
12. Saitta C. El timbre como factor estructurante. Altura-timbre-espacio. Cuaderno de estudio. 2016;5.
13. Miyara F. Mediciones acústicas basadas en software. Springer International Publishing AG; 2017.
14. Anderson J, Basso G, Cetta P, Cura M, Di Liscia OP, Dow R, Fessel P, Kendall GS, Liut M, Malham D, Pampin J. Música y espacio: ciencia, tecnología y estética. Univ. Nacional de Quilmes Ed.; 2009.
15. McLuhan E, Zingrone F. Essential McLuhan. Routledge; 1997 Jun 5.
16. Kassier R, Brookes T, Rumsey F. A simplified scene-based paradigm for use in spatial audio listener training applications. In117th AES Convention 2004 (p. 18).
17. Rodríguez, L. E. Construcción de mapas sonoros sobre música realizada en estéreo. En “Investigaciones sobre audio espacial y estética del arte sonoro” p. 7 – 46 Wolkowicz Editores
18. Rodríguez, L.E. Fuentes musicales como constelaciones de componentes dentro del espacio virtual. En “Avances del Audio en Latinoamérica 2022”, 2023. Proceedings of the AES LAC 2022 Conference. Audio Engineering Society Argentina, Instituto Terciario TAMABA.

FUNDING
CONICET – Internal Doctoral Scholarship

CONFLICT OF INTERESTS
None