Classification of Triadic Chord Inversions Using Kohonen Self-organizing Maps

Luis Felipe de Oliveira[1], Luis Guilherme Pereira Lima[2], André Luiz Gonçalves de Oliveira[3], Rael Bertarelli Gimenes Toffolo[1]

[1] Departamento de Música – Universidade Estadual de Maringá (UEM)
Maringá – PR - BR

[2] Departamento de Informática – Universidade Estadual de Maringá (UEM)
Maringá – PR – BR

[3] Departamento de Teoria e Prática da Educação – Universidade Estadual de Maringá (UEM)
Maringá – PR – BR

luisfol@bol.com.br, lglima@gmail.com, alguns@gmail.com, rael_gimenes@uol.com.br

Abstract. *In this paper we discuss the application of the Kohonen Self-organizing Maps to the classification of triadic chords in inversions and root positions. Our motivation started in the validation of Schönberg´s hypotheses of the harmonic features of each chord inversion. We employed the Kohonen network, which has been generally known as an optimum pattern classification tools in several areas, including music, to verify that hypothesis. The outcomes of our experiment refuse the Schönberg´s assumption in two aspects: structural and perceptual/functional.*

1. Schönberg's Hypothesis

In the *Harmonielehre*, Schönberg (1974) discuss the peculiar harmonic features and rules of the first and second triadic inversions at the relation of root position triads (53). The first inversion (6) does not require any kind of special treatment than the second inversion (64) does. The second inversion, as he argues, has an ambiguity constitution, being it related to its root position chord and to a chord a fifth above. This ambiguity has been lead to specific harmonic rules in the attempt to characterize the function of this chord. For instance, the tonic second inversion chord can be analyzed as I64 or V64, and in this case, the cadential 64, there are no specific voice leading procedures in bass. In Schönberg´s assumption there is a contrasting forces involved in the 64 chord due by the derived harmonic partial structure. There are harmonic partials that emphasize the bass note as fundamental (they match the bass harmonic series) and there are partials that attenuate the strength of the bass as fundamental (they do not match the bass harmonic series). In the first inversion chord, the harmonic partials that match the bass note (in this case, the third) are higher than the ones that match the fundamental. In the same time, at the second inversion chord, the harmonic partials that matches the bass note (in this case, the fifth) are closer than ones that matches the fundamental. One can resume the Schöenberg´s treatment of second inversion chord as moving the bass voice by step wise or staying on the same note, after and before this chord, producing an aural perception of a chord over a passing note, solving the ambiguity. For the sake of clarity, the above statement can be illustrated in the following table:

Table 1. Harmonic partials set in a C major root position and inverted chords (the played notes in capitals letters; the harmonic partials in lower case; the subscript letters are not considered by Schönberg; the notes that matches the bass in blue and the harmonic partials that contradict the bass in red).

Harmonic partials of triadic chords																			
Root position (53)																			
C			c		g		c		e	g	$_b$b								
	E			e		B			e	g#		b							
		G			g			D		G		b		d					
First inversion (6)																			
	E			e		B			e	g#		b		$_d$	$_e$				
		G			g			d		G		b		d					
			C				c			G			c		e			G	
Second inversion (64)																			
		G			g			d		G		b		d		$_f$	$_g$		
			C				c			G			c		e			G	
				E					e			b			e			g#	b

Analyzing the table and considering the contradictory harmonic partials, which are not considered by Schönberg, one can note that the root position presents problems in the same way that the second inversion does. We must stress that Schönberg do not note the harmonic sets to the root position chord - in this sense, his argument became fragile and naive. Otherwise, he neither evaluate this explanation to the minor tonalities, what seems to be a recurrent conduct of music theorist in general.

In Oliveira *et al* (2005), we demonstrated the fallacy of Schönberg´s argumentation for a single major tonality using a Kohonen Self-organizing Maps. The experiment results, demonstrated that the root position chords and its inversions were classified in near clustering neighborhood. If the Schönberg´s hypothesis were correct, the 64 chords should be placed in a fifth above chord neighborhood.

In this paper, our inquiry is motivated by the historical lack in music theory of a solid argumentation explaining the minor chords. The supposed perfection of major triad was explained by the relations of the harmonic partials, especially the six foremost ones[1]. If there are difficulties to justify the major chord "perfection" in this argument, in the minor chords, the argumentation problems about its constitution and tonal origins, will be more prominent. To escape from this traditional lack, present in harmony studies, we must verify the results of a SOM working on minor, major and diminished chords as input values. Our input patterns set consists in the 252 chords from the twelve major tonalities (21 chords for each tonality). The use of a SOM to classify chords is well known in the literature about cognitive (and computer assisted) musicology [LEMAN

[1] If the perfection is derived from the harmonic series, one can note that the seventh partial brings troubles to this description.

2000, 1995, 1991, 1990]. However, usually this sort of research does not consider any aspect of chord inversion. That is the main objective of our investigation. We expect that the clustering map will create the same categorization for the inverted chords for all major, minor and diminished triads. The results should appoint that the problem explanations based in the harmonic series, that justify the structure and the harmonic treatments of the inverted chords, are misleading for major, minor and diminished chords.

2. The Kohonen's Self-organizing Maps.

2.1. The Self-organizing Maps.

The main objective of a Kohonen Self-organizing Map (SOM) (1997) is to determine the mapping of a *n* dimensional input signals set into a bi-dimensional grid. This mapping occurs in an adaptative and topologically ordered fashion. The SOM architecture consists in two layers: the input layer of *n* dimensionality (*n* equal to the dimensionality of the inputs set) and the output layer characterized as a bi-dimensional grid of neurons. Each input neuron is fully connected with the bi-dimensional grid, where each of the connections is represented by an associated synaptic weight. The weight vector of a grid unit consists in the set of weight values of all connections.

The algorithm responsible to the map formation initializes the grid connections weight with small random values. We can stand out three processes when the network is initialized: competition; cooperation; synaptic adaptation. These processes are responsible to the formation of the map in an unsupervised learning method [HAYKIN 2001].

Competition:

When an input pattern is presented to the network, just one grid unit must be activated by that pattern. This activated unit is called the winner neuron (winner-take-all). A discriminant function provides the competition bases in this process.

Cooperation:

The winner neuron determines the center of the topological neighborhood region of laterally excited neurons, furnishing the bases of cooperation. The neighborhood area can be expressed as a rectangular field or a hexagonal one, as showed in the figure bellow.

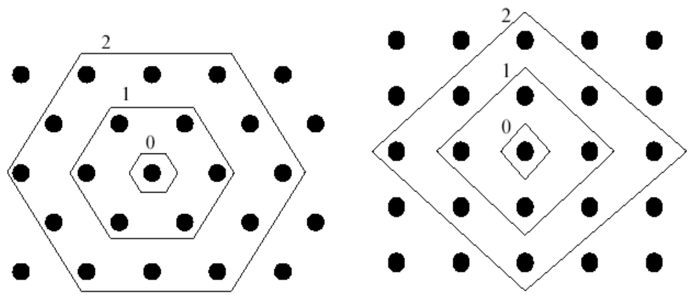

Figure 1. Rectangular and hexagonal neighborhood.

The lateral interaction among the winner neuron and its neighbors is represented by the neighborhood topological function. The maximal function value is reached when the distance among the winner neuron and its neighbors is equal to zero. Otherwise, the minimum function value occurs when the distance among the winner neuron and its neighbors tends to the infinity. The gaussian function satisfies the above necessities and is widely employed in SOM networks:

$$f_{j,i(x)} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right)$$

where the dj, i represents the lateral distance between the winner neuron i and the excited neuron j. The σ represents the effective width of neighborhood as showed in the Figure 1.

Synaptic adaptation

In this phase, the synaptic weight are adjusted to that the winner's weight vector approximates to the input vector x. Once the network is continuously feed with the input set, the algorithm produces a topological map ordination of the features, leading to similar values of the weight vectors for the adjacent neurons.

2.2. The SOM Algorithm

The SOM Algorithm could be summarized in five steps:

1. Initialization

Attribute random small values to the synaptic weights of each neuron. This step ensures that the map will have no previous organization at all.

2. Sampling

Present a random choose sample X to the network, from the input space with n dimensionality, being $X = [x_1, x_2, x_3, ..., x_n]^t$.

3. Matching

Find the winner neuron $i(X)$, with the weight vector $W = [w_1, w_2, w_3, ..., w_n]^t$, in the epoch t, closer to the vector X presented to the network, adopting the minimum Euclidian distance criterion:

$$i(X) = \arg\min_j \|X - W_j\|, j = 1, 2, ..., n$$

being n, the total number of grid neurons.

4. Updating

Adjust the synaptic weights to every grid neuron by the actualization formula

$$w_j(t+1) = w_j(t) + \eta(t) f_{j,i(x)}(t)(x(t) - w_j(t))$$

where $\eta(t)$ is the learning rate, $f_{j,i(x)}(t)$ is the neighborhood topological function surrounding the winner neuron. Both parameters are dynamically varied to ensure better results.

5. Repetition

Return to step 2 until no significant changes occur in the features map.

The competitive learning happens in the second step when the weight vectors are updated. To each input presented to the network only one neuron must be active in a specific instant. In this sense, the competitive learning has a priority to define the statistical features more outstanding for the classification of an input pattern set [HAYKIN 2001]. In the third step, the cooperative process among the grid neurons is determined by the neighborhood area of the winner unit ($f_{j,i(x)}$).

The network training consists in two distinct phases: the rough phase and the fine-tuning phase. The first is characterized by the topological ordination of the weight vectors. In this phase, the learning rate should be set in value close proximity to 0.1 and the neighborhood area of the winner neuron should take almost all neurons of the grid. During the rough phase the learning rate decreases smoothly until it reaches the value of 0.01. The second phase is necessary to achieve a fine-tuning ordination of the features map. To make the statistical precision as good as possible, the learning rate should be maintained closer to 0.01; it should not take zero to avoid a meta-stable state of the grid (a topological impairment). The neighborhood area must contain only the next neighbors of the winner unit, possibly reducing its area to one or zero in the fine-tuning phase.

3. Method

The network input dataset is a group of vectors with dimensionality n=63, being n the necessary number of pitch-classes to represent the sixth first harmonic partials of each chord note, without octave reduction. The value of the partials in each vector was set up in two fashions: decreasing Fibonacci scaling and an unique value of 1 for all the six partials. Both alternative scaling procedures have leaded the network to similar results [OLIVEIRA et al 2005], and we adopted the Fibonacci scaling for the experiments described in this paper. The SOM network employed has a total of 2500 units, as a square matrix of 50 x 50 cells.

To make the topological error as minor as possible, we adopted the optimum-learning rate of $\alpha 1=0.1$ and $\alpha 2=0.037$ for the rough and fine-tuning training epochs, respectively. The rough epoch consisted in 500 iterations and the fine-tuning in 1500 (the network training employed a learning rate power function in sequential input presentation mode). The neighborhood radius was started with 30 units (rough phase) and decreases to 1 (fine-tuning phase).

4. Results and Discussion

During the network training, the quantization error was reduced to zero, as shown in figure 2, in the fine-tuning training phase. After the training period we can obtain the topological representation of the network response to the input set, as figure 3 demonstrates.

The topological map shows the clusters of each pattern of the input set. In a very first observation one can see that the map demonstrates that similar chords are grouped together. This fact proves that the outcome is coherent over the network behavior. Similar chords have common values in the input vector, so it’s natural that they are placed in near positions of the topological map. For instance, root position chords were classified as very similar categories than the inverted ones. One can ask why a chord of C6 is classified closer to C53 than Em53 if it has two common notes with both chords. Could be argued that the fact it was classified near to C53 than Em53 because the octave interval has common values in the input vector but the minor second (obviously!). In this sense, the octave similarity is due to the equivalence of some input vector values.

Regarding Schönberg’s argumentations about the 64 chords, the network results should put this sort of chord far from the root position ones. The result, as in the figures bellow, clearly dismisses such an argumentation. The argument of the author came from the fact that there are harmonic partials that stress the bass note and others that emphasize the fundamental note. One can take this statement as attempt to include perceptual properties in the scope of the explanation about chord similarities. On the other hand, one can take that Schönberg’s assertion is an attempt to include functional statements into the justification. But, besides his statements, the SOM network placed the 64 chord generally in the same cluster that the 53 chord, as being the same category. The topological map grouped the chords and its inversions in the same neighborhood in 97.22% of the cases (see figure 5). The isolated chords, beside the fact they are placed distantly from its similes, they are distant from the neighbor chords as should be noted in a three-dimensional representation of the topological map (figure 4). The presence of the hills on this map indicates the growing of the Euclidian distance to the non-similar neighbors. This statement reveals that the SOM network classified those chords in localizations far away from their similes, but categorized they as being different from the other chords in the same topological region.

Recalling that the 64 inversion was classified as the same category of the root position chord, the overall outcome corroborates the Enlightenment theorist Rameau (1971), in his argumentation of the chords inversions. Rameau did come to his results inside a deep reflection and sometimes an unusual mathematics over the monochord[2]. Our results came from the proximal values of the input vectors in octave interval (see table 2).

Table 2. Input vectors of a note C and octave above C. The boldface items show the coincident partials.

C	C#	D	D#	E	F	F#	G	G#	A	A#	B	**C**	C#	D	D#	E	F	F#	G	G#	A	A#	B	**C**	C#	D	D#	E	F	F#	**G**	G#
1	0	0	0	0	0	0	0	0	0	0	0	**0.8**	0	0	0	0	0	0	0.5	0	0	0	0	**0.3**	0	0	0	0.2	0	0	**0.1**	0
0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	**0.8**	0	0	0	0	0	0	**0.5**	0

A	A#	B	C	C#	D	D#	E	F	F#	G	G#	A	A#	B	C	C#	D	D#
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.3	0	0	0	0.2	0	0	0.1	0	0	0	0	0	0	0	0

Based on the results of the SOM network, we can say that Schönberg's affirmations and our refutations are founded on two points: structural and perceptual/functional. Firstly, we will analyze the structural point in the bases of three aspects: the temperament; the consideration of only the sixth firsts harmonic partials; the dynamic behavior of harmonic partials. Secondly, the argumentation about the perceptual/functional considerations of inverted chords should be took in a different way that Schönberg considers it, if he does at all (in *Harmonielehre e* specially[3]).

In the structural consideration, the non-tempered harmonic series used in the Schönberg's statement, does not seems to be the better way to explain the emphasis on the bass notes by the coincident harmonic partials generated by the other chords notes. To evaluate his hypothesis we adopted the same representation in the construction of the inputs patterns, i.e. non-octave reduced pitch class. In this sense, we had model the experiment using the same presuppositions provided by Schönberg, even so his hypothesis was not confirmed. Other issue that must be considered is the use of only the six first partials. Schönberg uses only this set of partials to avoid the problems that the seventh partial brings to the scene[4]. The non-consideration of the seventh partial is misleading in the sense that one uses just the corroborative aspects of a natural phenomenon for the sake of argument. We can suppose that this fact is justified by the necessity to establish a natural foundation to the tonal system. This natural foundation is as artificial as any abstract and arbitrary possible syntactic system (twelve-tone scale or

---

[2] For furthers stances see Rameau (1971) and specially the translator's introduction.

[3] Even in Structural Functions of Harmony, Schönberg does not talk about chord inversions.

[4] Rameau, in his treatise, spares the very same problem skipping the seventh partial from his explanations (1971, p.7).

microtonal scale etc) because they did not take the nature phenomena as whole coherent instance but just some parts of it. The third point in the structural aspects concerns the dynamic behavior of harmonic series (in nature). Beside the above considerations the harmonic series is not a static and stable set of frequencies over the time independent from its physical instantiation. If the explanation for the inverted chords bears on the harmonic partials, a chord structure played in some of the musical instruments or electronic devices should be very confusing for a listener. For instance, a second inversion chord played by the clarinetists does not presents the same set of harmonics as an idealized harmonic series. Consequentially, the explanation of chord inversions would not apply to such a physical instantiation. Could the relation of a chord and its inversions be accounted in the perspective of an idealized and incomplete harmonic series? The present experiment does not take this fact into consideration (our goal was just test the Schönberg's hypothesis), but further developments we are designing operate over a non-tempered harmonic series with spectral envelope behavior extracted from various musical instruments.

To conclude our divagation we should state some perceptual (and functional) considerations. Perhaps, the better justification for the closer relation of a 64 chord with a 53 chord a fifth above came from a perceptual *cliché* and some counterpoint procedures, as the *retardo* or suspension (the dissonance should ever descend like all the things fall to the ground, as stated by Fux (1965) in the XVIII century). The perceptual justification does not necessarily require an acoustic correlation as Schönberg tried to do, because sometimes it is more a subject of cultural stereotyped expectations, that are meaningful just to a part of the human beings, and is not absolutely a property of nature. Being that, perhaps the some music theorists, in an attempt to justify the naturalness of some abstract system, attaches a physicalist reasoning to some properties and procedures coming from the level of language and art, that are not just a physical phenomena. Finally, we can recall that the Schönberg's physical hypothesis to the second inversion chord, based on a static and idealized harmonic series, was not confirmed by the experiment described in this paper. In further researches we intend to examine the limits of the relation of tonal functionality and structural acoustics aspects of triadic chords, using a representational system closer to the acoustic phenomena and capable to account for harmonic progressions.

4.1. Graphic Results

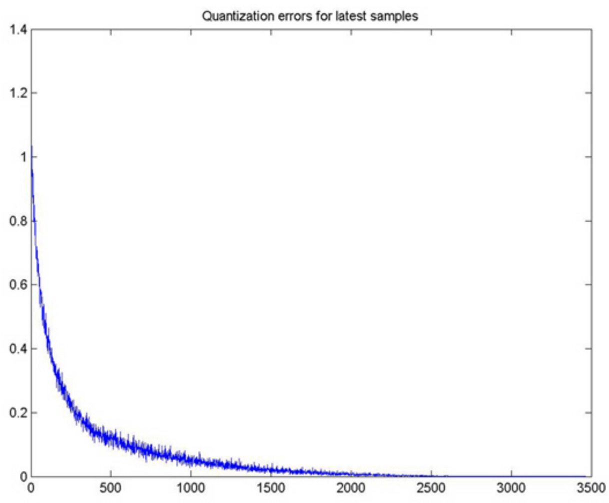

Figure 2. Quantization error.

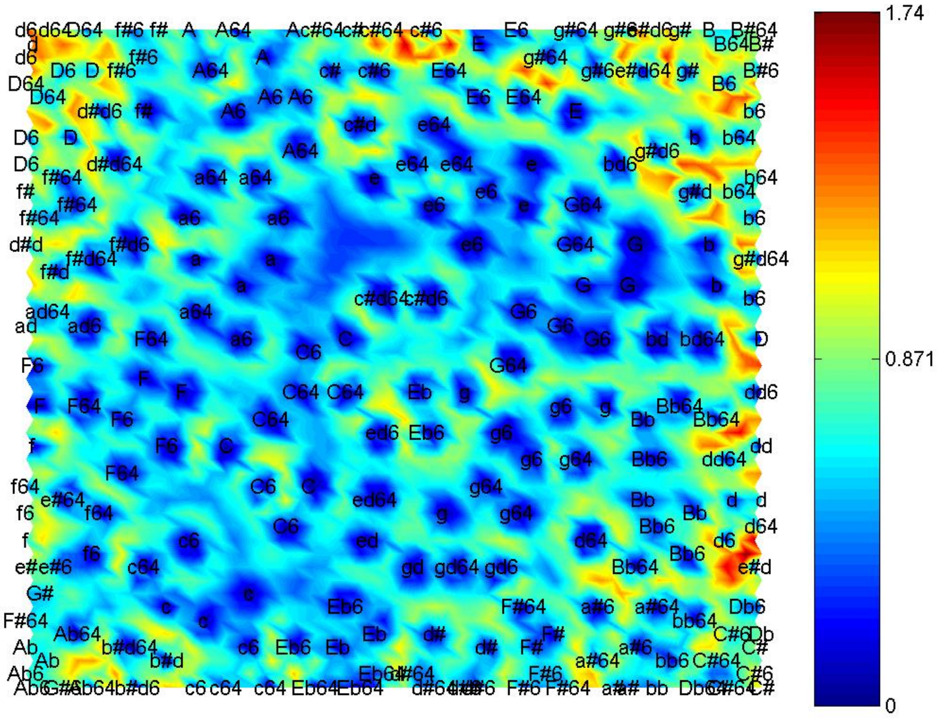

Figure 3. Topological Map. The color bar indicates the Euclidian distance among clusters.

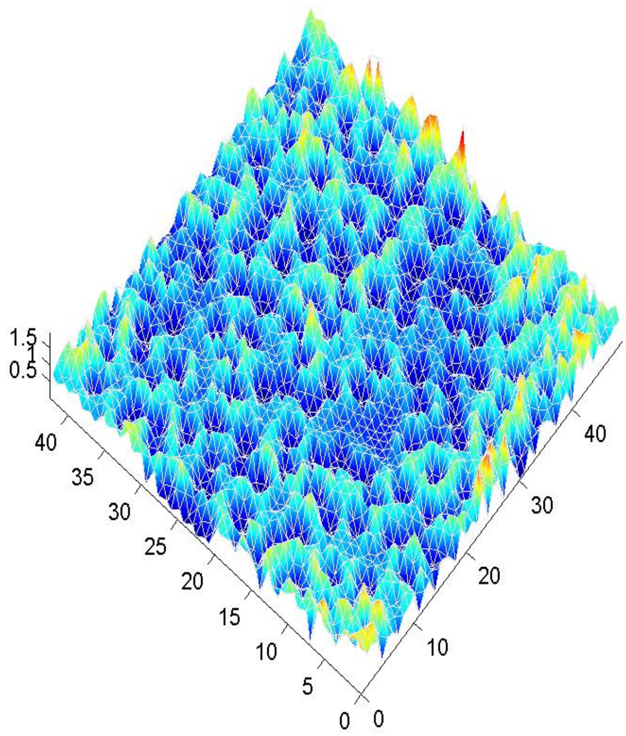

Figure 4. The three-dimensional representation of the topological map.

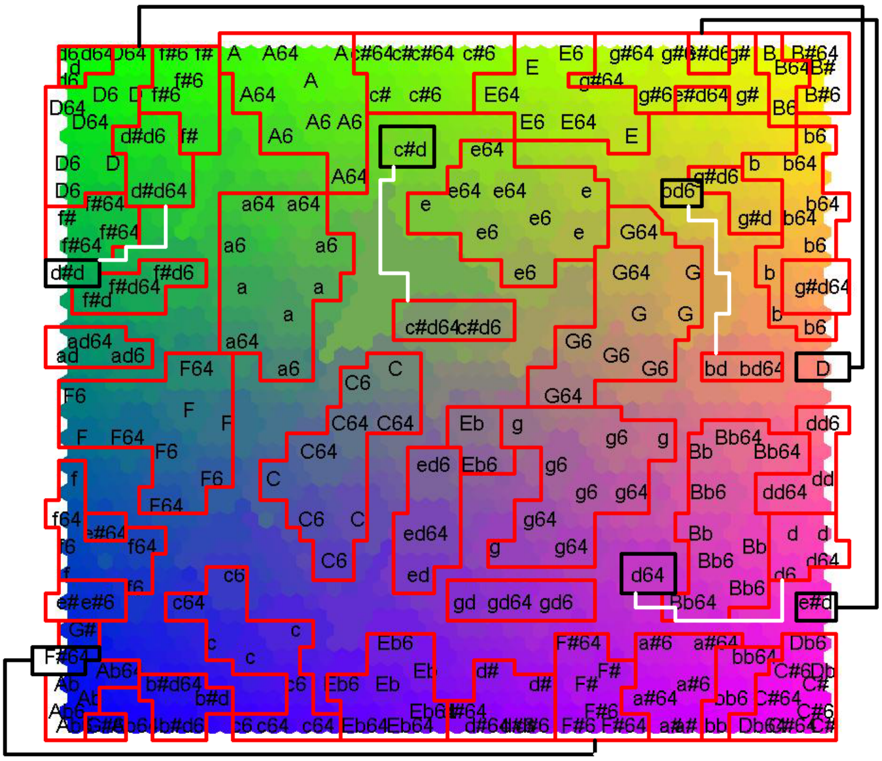

Figure 5. The grouping of the chords and its inversions. Encapsulated in black are the isolated chords, linked with its similes.

References

Arbib, Michael A. (1998). The Handbook of Brain Theory and Neural Networks. Cambridge, MA: The MIT Press.

Fux, Johann J. (1965). The Study of Counterpoint From Johann Joseph Fux’s Gradus ad Parnassum. New York: W. W. Norton & Company.

Leman, Marc (2000). An Auditory Model of the Role of Short-Term Memory in Probe-Tone Ratings. *Music Perception*, Vol.17(4), 481-509.

Leman, Marc (1995). A Model of Retroactive Tone-Center Perception. *Music Perception*, Vol.12(4), 439-471.

Leman, Marc (1991). The Ontogenesis of Tonal Semantics: results of a computer study. In: P. Todd & G. Loy (Eds.), *Music and Connectionism*. Cambridge, MA: The MIT Press.

Leman, Marc (1990). Emergent Properties of Tonality Functions by Self-organization. *Interface*, Vol.19(2-3), 85-106.

Haykin, Simon (2001). Redes Neurais: princípios e prática. Porto Alegre: Bookman.

Kohonen, Teuvo (1997). Self-organizing Maps. Berlin: Springer-Verlag.

Oliveira, Luis F, Lima, Luis G. P., Oliveira, André L. G. and Toffolo, Rael B. G. (2005). Classificações de Acordes Invertidos por Mapas Auto-organizados de Kohonen, In: M. Dottori , B. Ilari & R. C. Souza (Eds.), *Anais do Primeiro Simpósio Internacional de Cognição e Artes Musicais*. Curitiba: UFPR.

Rameau, Jean-Philippe (1971). Treatise onHarmony. New York: Dover Publications.

Schönberg, Arnold (1974). Armonia. Madrid: Real Musical.