The Influence of Music on the Emotional Interpretation of Visual Contexts

Designing Interactive Multimedia Tools for Psychological Research

Fernando Bravo

University of Cambridge. Centre for Music and Science.
nanobravo@fulbrightmail.org

Abstract.
From a cognitive standpoint, the analysis of music in audiovisual contexts presents a helpful field in which to explore the links between musical structure and emotional response.

This work emerges from an empirical study that shows strong evidence in support of the effect of tonal dissonance level on interpretations regarding the emotional content of visual information.

From this starting point it progresses toward the design of interactive multimedia tools aimed at investigating the various ways in which music may shape the semantic processing of visual contexts. A pilot experiment (work in progress) using these tools to study the emotional effects of sensory dissonance is briefly described.

Keywords: Music, emotions, film-music, interactive multimedia, algorithmic composition, dissonance, tonal tension, interval vector, Max/MSP/Jitter.

1 Introduction and Background

Although research in music cognition has been growing steadily during the past four decades, we still lack a significant body of empirical studies concerning the higher levels of musical response, including the emotional and aesthetic aspects. From a cognitive standpoint, the analysis of music in audiovisual contexts presents a helpful field in which to explore the affective and connotative aspects of musical information [1, 2, 3].

This paper describes work in progress to investigate the influence of tonal dissonance on the emotional interpretation of visual information.

The objectives of this paper are:

- To report on a formal experiment showing the effect of tonal dissonance on interpretations regarding the emotional content of an animated short film (Section 2).
- To describe a series of interactive multimedia tools designed to investigate the various ways in which music may shape the semantic processing of visual contexts (Section 3).

9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)
19-22 June 2012, Queen Mary University of London

- To show an example of how these tools could be used in experimental cognition research. In this example, I employ stochastically generated music to empirically study the links between sensory dissonance and emotional responses to music in a strictly controlled audiovisual setting (Section 3.2).

Consonance and dissonance refer to specific qualities an interval can posses [4]. Tonal and sensory dissonance are sometimes used as equivalent concepts. However, as Krumhansl [5] has expressed, these two notions have different shades of meaning. Sensory dissonance designates, first of all, a psychoacoustic sensory property associated with the presence/absence of interaction between the harmonic spectra of two pitches [6]. Tonal dissonance includes sensory dissonance but it also captures a more cognitive or conceptual meaning beyond psychoacoustic effects that is typically expressed with terms such as tension or instability. The term "tonal dissonance", as employed here, refers both to sensory and cognitive dissonance.

Meyer proposed a theory of meaning and emotion in music [7]. According to his assumptions the confirmation, violation or suspension of musical expectations elicits emotions in the listener. Following this theory, researchers found association between specific musical structures, precise neural mechanisms and certain neurophysiological reactions that are strongly connected with emotions. In addition, studies focusing on the perception of tonal dissonance have shown that unexpected chords and increments in dissonance have strong effects on perceived tension [5, 8, 9, 10], which has been linked to emotional experience during music listening [11].

Tonal dissonance can be described by a number of variables [9], which have been already historically studied by music theorists and scientists: the tonal function of chords inside a musical context [12, 13, 14, 15, 16], their acoustic or sensory consonance [6, 17], and melodic organization, usually referred as "horizontal motion" [18].

Cognitive approaches usually emphasize the importance of melodic organization and tonal function while sensory-perceptual theories tend to focus on psychoacoustical aspects. In this paper, I use the term 'tonal dissonance' as a synonym for 'tonal tension', to refer to the effects of tonal function, sensory dissonance and horizontal motion on perceived musical tension.

2 Experimental Investigation

This paper emerges from a formal experiment entitled "The influence of tonal dissonance on emotional responses to film" [19]. The main experimental hypothesis predicted that, within the same film sequence (visual context), different musical settings, in terms of tonal dissonance, would systematically elicit different interpretations and expectations about the emotional content of the same movie scene.

2.1 Experimental Design

This experiment was aimed at addressing the particular emotional effect of tonal dissonance induced by chord changes, controlling for other elements within musical structure such as tempo, intensity, rhythm, timbre (instrumentation), etc. This was achieved by working with a precise experimental design, also used by Blood *et al.* in their neuroscientific research (which investigated the cerebral activations elicited by tonal dissonance) [20]. It is important to note that this study sets aside other kinds of musical tension. Empirical evidence has shown that musical tension can be induced by many factors, such as rhythm, dynamics, tempo, gesture, textural density and tone timbre [25, 26, 27]. This work focuses on musical tension induced by tonal dissonance in the specific sense of tension created by melodic and harmonic motion.

A choral piece, specifically composed for the experiment, was made to sound more or less consonant or dissonant by modifying its harmonic structure, producing two otherwise-identical versions of the same music passage. These two contrasting conditions, in terms of tonal dissonance, were used as background music for the same passage of an animated short film ("Man with pendulous arms" - 1997, directed by Laurent Gorgiard).

Table 1. Cross-classification of music condition and response variable (number of participants and percentage of participants within condition)

The character	*feels confident*		*is scared*	
consonant	37	61.7%	23	38.3%
dissonant	25	41.7%	35	58.3%
The mood of the story is	*nostalgic*		*sinister*	
consonant	58	96.7%	2	3.3%
dissonant	26	43.3%	34	56.7%
The character is trying	*to create something*		*to destroy something*	
consonant	45	75%	15	25%
dissonant	27	45%	33	55%
The character	*is a fantasy character*		*is monstrous*	
consonant	53	88.3%	7	11.7%
dissonant	39	65%	21	35%
Genre of the short film	*Drama*		*Horror*	
consonant	59	98.3%	1	1.7%
dissonant	42	70%	18	30%
The character	*is alienated*		*is sad*	
consonant	11	18.3%	49	81.7%
dissonant	35	58.3%	25	41.7%
Character's actions	*directed by his own will*		*external influence*	
consonant	50	83.3%	10	16.7%
dissonant	35	58.3%	25	41.7%
The end of the short film	*will probably be hopeful*		*will probably be tragic*	
consonant	41	68.3%	19	31.7%
dissonant	29	48.3%	31	51.7%
The character is trying	*to protect himself*		*to search something*	
consonant	47	78.3%	13	21.7%
dissonant	44	73.3%	16	26.7%

A total of 120 healthy volunteers with normal hearing took part in this experiment. The participants were randomly sampled from students at Argentine Catholic University. Two independent samples were used (60 participants each). The subjects were randomly assigned to two groups, one of which saw an animated short film with the "consonant music" condition and the other saw the same film with the "dissonant music" condition. At the end participants were asked to answer a survey about their associations and expectations towards the main character and the overall story of the film. The survey used 9 single-selection questions, asking participants to choose only one item from two items given. Table 1 shows participants' answers within each music condition.

2.2 Experimental Results

Eight out of nine response variables were found associated with the explanatory variable (tonal dissonance level). The variable related to the character's objective (at the bottom of Table 1) was the only variable that did not reach significant association with tonal dissonance level.

For the eight response variables where association was found, two ways to summarize the strength of the association are presented: the *difference of proportions*, forming confidence intervals to measure the strength of the association in the population, and the *odds ratio* (Table 2).

When measuring the strength of the association, variables related to the mood in the story, the emotional state of the character and the interpreted genre of the short film were found to have the strongest association with dissonance level in background music (grey cells).

Table 2. $\chi 2$, Difference of proportions and Odds Ratio

		Differ. of proportions		Odds Ratio	Odds	
Variable	$\chi 2$(*p value*)	$p1$-$p2$	95% CI	OR	Con	Dis
Intentions (create/destroy)	11.2(<.01)	0.3	[.133, .467]	3.667	3	0.8
Feeling (confident/scared)	4.80(<.05)	0.2	[.025, .035]	2.252	1.6	0.7
Mood (nostalgic/sinister)	40.6(<.01)	0.53	[.401, .667]	37.92	29	0.7
Emot.state (sad/alienated)	20.3(<.01)	0.4	[.241, .559]	6.236	4.4	0.7
Actions (own will/external)	9.07(<.01)	0.25	[.094, .406]	3.571	5	1.4
Class (fantasy/monstruous)	9.13(<.01)	0.23	[.087, .379]	4.077	7.5	1.8
Genre (drama/horror)	18.0(<.01)	0.28	[.163, .403]	25.28	59	2.3
Ending (hopeful/tragic)	4.93(<.05)	0.2	[.027, .373]	2.307	2.1	0.9

For example, Table 2 shows that there was a rise of 0.4 in the proportion that interpreted the emotional state of the character as sad among participants who saw the film with consonant music. Also, we may infer, with 95% confidence, that $p1$ (the proportion of people seeing the film with consonant music and interpreting the character's emotional state as sad) may be as much as between [0.241, 0.559] larger than $p2$ (the proportion of people seeing the film with dissonant music and interpreting the character's emotional state as sad).

In addition, from Tables 1 and 2, we observe that for the consonant music condition the proportion of people who interpreted the character's emotional state as sad equals 49 / 11 = 4.4545. The value of 4.45 means that, for participants who saw the film with consonant music, there were 4.45 participants who interpreted the character's emotional state as sad, for every 1 person in the dissonant condition. On the other hand, for the dissonant music condition the proportion of people who interpreted the character's emotional state as sad equals 25 / 35 = 0.7143. Equivalently, since 35 / 25 = 1 / 0.7143 = 1.4, this means that there were 1.4 participants in dissonant condition who interpreted the character's emotional state as alienated for every 1 person in consonant condition. For the consonant music condition, the odds of interpreting the character's emotional state as sad were about 6.2 times the odds of the same interpretation for the dissonant music condition.

The results of this experiment offer strong evidence in support of the effect of tonal dissonance level (in film music) on interpretations regarding the emotional content of visual information.

2.3 Discussion of Experimental Results

The empirical research described supports and confirms previous research on mood congruency effects [1], and can be interpreted within Annabel Cohen's Congruence-Associationist framework of the mental representation of multimedia [2, 3].

In this work, tonal dissonance level was experimentally isolated in order to analyze a particular feature within the multiple musical structures that may elicit musical emotions. As pointed in section 2.1., this study was focused on musical tension induced by chord changes. Other important factors that contribute to the building and release of musical tension, such as timbre, dynamics, textural density, etc., which were controlled in the experiment, are not examined in the present discussion.

Results revealed that the background music significantly biased the affective impact of the short film. Generally, the consonant music condition guided participants toward positive emotional judgments, while dissonant music guided participants toward negative judgments. In addition, the dissonant background music seems to have rendered the interpretation more ambiguous when compared to the higher percentages for the positive judgments in the consonant condition. However, additional research is needed to further examine this hypothesis since the present experiment did not include a visual alone condition, which would be necessary to control for the effects of visual content by itself.

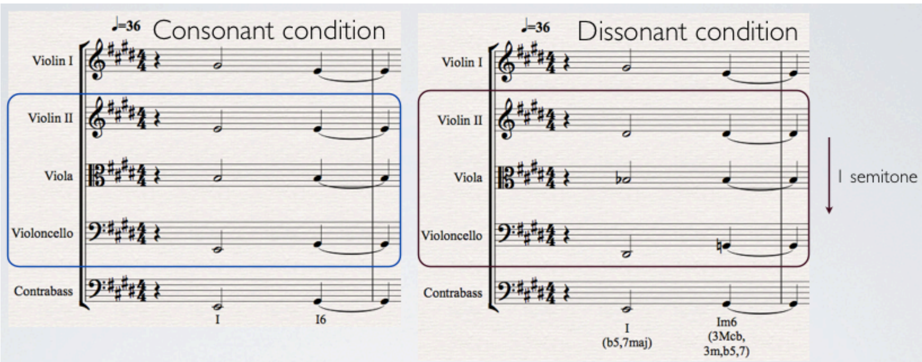
Music theory provides technical descriptions of how styles organize musical sounds and offers insights about musical structures that might underlie listeners' interpretations. Within the general perspective of post-tonal music theory, Allen Forte has introduced the notion of interval-class content [21]. This concept, widely used in the analysis of atonal twentieth-century music, offers an interesting approach to qualifying sonorities. A pitch interval is simply the distance between two pitches, measured by the number of semitones. The ordered pitch intervals (ascending or descending) focus attention on the contour of the line. The unordered pitch intervals ignore direction of motion and concentrate entirely on the spaces between the pitches. An unordered pitch-class interval is the distance between two pitch classes, and it is also called interval class [28]. Because of octave equivalence, compound intervals (intervals larger than an octave) are considered equivalent to their complements in mod 12. In addition, pitch-class intervals larger than six are considered equivalent to their complements in mod 12. The number of interval classes a sonority contains depends on the number of distinct pitch classes in the sonority. For any given sonority, we can summarize the interval content in scoreboard fashion by indicating, in the appropriate column, the number of occurrences of each of the six interval classes (occurrences of interval class 0, which will always be equal to the number of pitch classes in the sonority, are not included). Such scoreboard conveys the essential sound of a sonority.

Table 3 summarizes interval class content for the first measure of the experimental transformation used in this study to create contrasting conditions (Figure 1). The comparative dissonant condition was obtained by lowering, by a semitone, the second violin, viola and violoncello lines, while keeping the other instruments in their original position (at their original pitch). Thus, the level of dissonance was uniform throughout a given version. The analysis, therefore, can generally represent the comparative level of dissonance throughout.

Table 3. Interval Content of the two music conditions

Consonant condition - Interval Class content						
Interval Class	1	2	3	4	5	6
No. of occurrences	0	0	3	6	5	0
Dissonant condition - Interval Class content						
Interval Class	1	2	3	4	5	6
No. of occurrences	4	2	2	3	6	2

Fig. 1. Score of the consonant and dissonant music conditions (first measure)



The consonant condition is primarily governed by collections of intervals considered to be consonant (thirds, fourths and fifths). In contrast, the use of dissonant intervals in the dissonant version (major second, minor second and tritone) has a very specific emotional effect that is reflected in the participants' interpretations.

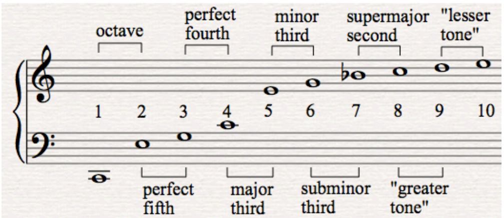
3 Future Work: Interactive Multimedia Tools for Experimental Research on Interval Content and Musical Emotions

The described experiment stimulated the investigation of tonal [15] and post-tonal [21] interval theory, and the parallel design of interactive multimedia tools to empirically analyze the effects of interval content on musical emotions.

3.1 Background Elements

Within the tonal perspective, Paul Hindemith's work is especially noteworthy [15]. According to Hindemith, the overtone series system (see Figure 2) gives a complete proof of the natural basis of tonal relations. In general, as new intervals are introduced, the stability decreases and the two tones involved are considered more distant in their relation. All music theories have a general agreement on this model.

Fig. 2. Overtone series with intervals labeled



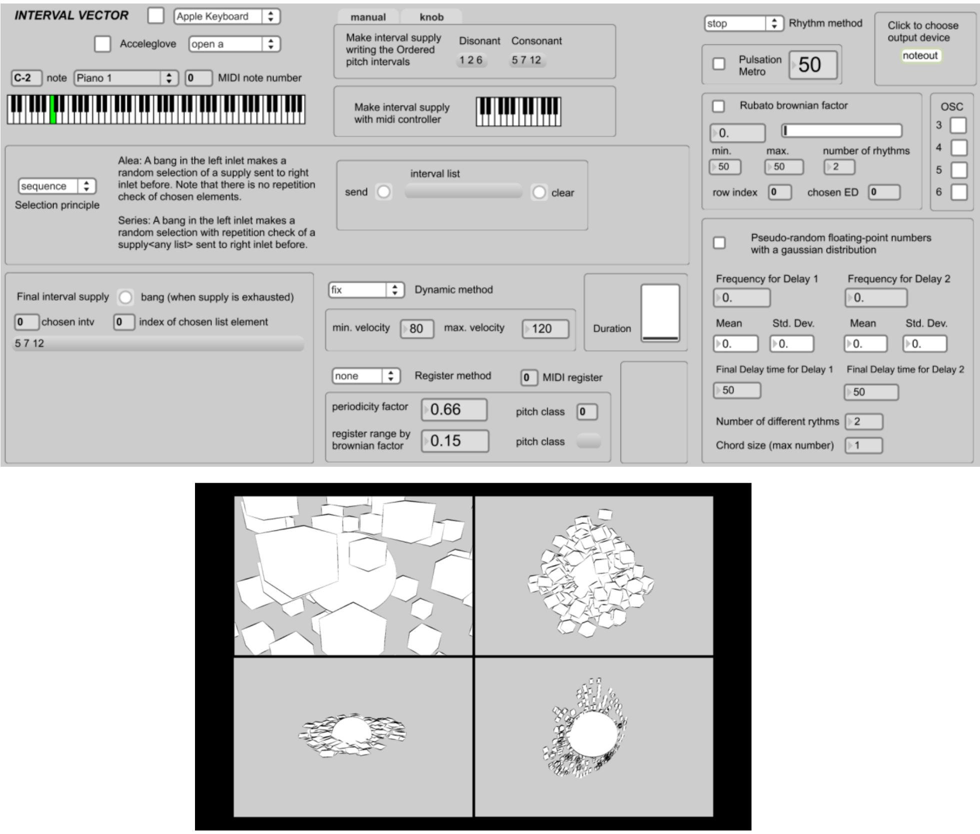
This theory has several links with the concept of sensory dissonance as studied in the psycho-acoustic literature [6, 22, 23, 24]. According to this model, the most consonant intervals would be the ones that could be expressed with simple frequency ratios, which has been supported by psychological study. Intervals such as the unison (1:1), the octave (2:1), perfect fifth (3:2), and perfect fourth (4:3) are regarded as the most consonant. Intermediate in consonance are the major third (5:4), minor third (6:5), major sixth (5:3), and minor sixth (8:5). The most acoustically dissonant intervals (composed of frequencies the ratio between which is not simple) are the major second (9:8), minor second (16:15), major seventh (15:8), minor seventh (16:9), and the tritone (45:32).

From the perspective of atonal theory, Allen Forte's work provided a general theoretical framework from where to start the exploration of intervals in a new way, a way that was intimately concerned with the idea of sonority [21]. He explained that different types of sonorities could be generally defined by listing their constituent intervals. In the previous experiment, I showed how the two music conditions could be described in terms of interval content. Forte introduced the basic concept of "interval vector" to analyze the properties of pitch class sets and the interactions of the components of a set in terms of intervals. An interval vector is an array that expresses the intervallic content of a set. It has six digits, with each digit standing for the number of times an interval class appears in the set [21]. According to Forte, such interval vector conveys the essential sound (color, quality) of a sonority.

3.2 'Intermedia Patch'. Cross-modal Research on Intervals and Visuals

The interactive multimedia tools presented in this section, called 'Intermedia patch', were built to explore the interval vector theory in a practical and strictly controlled setting, in order to experimentally study the links between sonority and emotional response. The patch works with an initial supply of intervals and allows to experiment with different algorithmic composition techniques, allowing a detailed control over many coincident variables such as loudness, rhythm, timbre, melody, intensity and instrumentation.

Fig. 3. Intermedia patch built with Cycling'74 Max/MSP/Jitter (top) that allows to simultaneously work with images created with Maxon Cinema 4D software (bottom)



The patch not only provides a programming environment for analyzing different types of sonorities based on interval selection (Figure 3 top), it also allows to simultaneously work with images (Figure 3 bottom), enabling the study of mood congruency effects between sound and visuals.

The tool is currently being tested in a pilot study (in progress), which employs the patch for the creation of sound stimuli. In this experiment I opted to analyze the emotional reactions induced by interval content. Participants are asked to see a short animation created for this study, with stochastically generated background music.

Participants are randomly assigned to three independent groups; one control group sees the animation without music, a second group sees the animation with a consonant interval content as background music (interval set: 5-7-12, all perfect consonances), and a third group sees the same animation with a dissonant interval content (1-2-6, all dissonances). Immediately after viewing the clip, participants are asked to complete a series of bipolar adjective ratings representing the three connotative dimensions: activity, potency and valence.

The question posed in this study is whether two contrasting examples of background music, in terms of interval content, can selectively bias observers' emotional interpretation of visual information. People who have internalized the Western tonal music conventions normally respond to certain sonorities in a specific manner. The main experimental hypothesis predicts that, in particular, the valence dimension should differ significantly under these two conditions. Positive results would confirm mood congruency effects induced exclusively by interval content (surface or sensory consonance).

4 Conclusions

The empirical research included in this paper supports and confirms previous studies that have examined, from a cognitive perspective, the role of music on the interpretation of a film or a video presentation [1, 2, 3]. The results offer strong evidence in support of the effect of tonal dissonance level on interpretations regarding the emotional content of visual information. Moreover, it gives insights to the richness and potentiality of the aural "palette", since extensive effects on the emotional interpretation of visual contexts may be directed by the manipulation of a single musical structure feature (tonal dissonance).

Studies such as this demonstrate associations between aspects of musical structure and musical meaning, which then becomes automatically attached to the visual content or implied narrative that is in the focus of the spectator's attention.

The positive results of this study indicate that further research that systematically examines the multiple and subtle ways in which music performs elaborative functions in the comprehension of visual contexts should be pursued. The interactive multimedia tools introduced in section 3 are aimed at exploring this path. These tools incorporate a variety of potential variables in both musical sound and transformations of the visual stimuli for experimental purposes, providing a foundation on which future research could build.

Acknowledgments. Thanks to Prof. Ian Cross, Prof. Sarah Hawkins and to all the researchers at the Centre for Music and Science (University of Cambridge). Thanks to Dr. Christopher Hopkins, Prof. Anson Call and Prof. Steve Herrnstadt for their constant support. Thank you to the anonymous reviewers for their suggestions that improved the paper considerably. This work was conducted at the University of Cambridge and is supported by a Queens' College Walker Studentship.

References

1. Boltz, M. G. (2001). Musical Soundtracks as a Schematic Influence on the Cognitive Processing of Filmed Events. *Music Perception,* 18, 427-454.
2. Cohen, A. J. (2001). Music as a source of emotion in film. In Juslin P. & Sloboda, J. (Eds.). Music and emotion. (pp.249-272). Oxford: Oxford University Press.
3. Cohen, A. J. (2005). How music influences the interpretation of film and Video: Approaches from experimental psychology. In R.A. Kendall & R. W. Savage (Eds.). Selected Reports in Ethnomusicology: Perspectives in Systematic Musicology, 12, 15-36.
4. Bharucha, J.J. (1984). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology*, 16, 485-518.
5. Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception*, 24, 329-366.
6. Helmholtz, H. von (1954). *On the Sensation of Tone as a Physiological Basis for the Theory of Music*. New York: Dover. (Original German work published 1863).
7. Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
8. Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, 58, 124-141.
9. Bigand, E., & Parncutt, R. (1999). Perception of musical tension in long chord sequences. *Psychological Research*. 62, 237-254.
10. Krumhansl, C.L. (1996). A perceptual analysis of Mozart's Piano Sonata, K. 282: Segmentation, tension and musical ideas. *Music Perception*, 13, 401-432.
11. Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18(8), 1380-1393.
12. Riemann, H. (1896). *Harmony simplified* (H. Bewerung, Trans.). London: Augener. (Original work published 1893).
13. Koechlin, C. (1930). *Traité de l'harmonie*. Paris: Max Eschig.
14. Schenker, H. (1979). *Free Composition* (E. Oster, Trans.). New York: Longman. (Original work published 1935).
15. Hindemith, P. (1942). *Unterweisung im Tonsatz*, 3 vols. (Mainz: Schott, 1937–70) [English edition, as *The Craft of Musical Composition*, vol. 1: Theoretical Part, trans. by Arthur Mendel (New York: Associated Music Publishers; London: Schott, 1942)].
16. Costère, E. (1954). *Lois et styles des harmonies musicales*. Paris: Presses Universitaires de France.
17. Rameau, J.P. (1971). *Treatise of Harmony* (P. Gosset, Trans.). New York: Dover. (Original work published 1722).

18. Ansermet, E. (1961). *Les fondements de la musique dans la conscience humaine*. Neuchâtel: Delachaux et Niestle.
19. Bravo, F. (2011). *The influence of music on the emotional interpretation of visual contexts* by *Bravo, Fernando.* Master's Thesis*,* Iowa State University, United States. AAT 1494771.
20. Blood, A.J., R.J. Zatorre, P. Bermudez & A.C. Evans. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci*. 2: 382–387.
21. Forte, A. (1973). *The Structure of Atonal Music*. Yale University Press.
22. Plomp, R. & Levelt, W.J.M. (1965). Tonal consonance and the critical bandwidth. *Journal of the Acoustical Society of America, 38*, 548-560.
23. Vos, J. & van Vianen, B.G. (1984). Thresholds for discrimination between pure and tempered intervals: The relevance of nearly coinciding harmonics. *Journal of the Acoustical Society of America, 77*, 176-187.
24. DeWitt, L.A. & Crowder, R.G. (1987). Tonal fusion of consonant musical intervals. *Perception & Psychophysics, 41*, 73-84.
25. Barthet, M., Depalle, P., Kronland-Martinet, R., Ystad, S., (2010). From clarinet control to timbre perception. *Acta Acustica united with Acustica*, 96, 678-689.
26. Barthet, M., Depalle, P., Kronland-Martinet, R., Ystad, S., (2010). Acoustical correlates of timbre and expressiveness in clarinet performance. *Music Perception*, 28, 135-153.
27. Paraskeva, S., McAdams, S. (1997). Influence of timbre, presence/absence of tonal hierarchy and musical training on the perception of tension/relaxation schemas of musical phrases. *Proceedings of the 1997 International Computer Music Conference, Thessaloniki*, pp.438-441.
28. Strauss, J.N. (1990). *Introduction to Post-Tonal Theory*. Prentice-Hall, Cliffs.