VOX PHENOMENA

A PSYCHO-PHILOSOPHICAL
INVESTIGATION OF THE PERCEPTION OF EMOTIONAL MEANING IN THE
PERFORMANCE OF SOLO SINGING (19TH CENTURY GERMAN LIED
REPERTOIRE)

ANTÓNIO G. SALGADO

A thesis presented to the Music Department
of the University of Sheffield in candidacy
for the degree of Doctor of Philosophy

February 2003

ABSTRACT

The research in this thesis is a contribution to the study of expression in music performance. It is primarily concerned with the study of emotion expressiveness in the performance of singing. This thesis focuses on the Western Art Music Canon and with solo singing, in particular, (within the Nineteenth Century German Lied repertoire) involving verbal and/or dramatic action. Though always considered within its musical context, the emotionally expressive character of singing with the words and/or a dramatic context, usually with an explicit narrative, is taken together and considered as a complex yet integrated whole. Two different approaches are adopted:

- Investigating the production of the facial movements and vocal sounds directly involved within the expression of emotion in singing.

- Exploring how the performed elements are perceived, recognised and experienced by the audience.

The movements and sounds involved in the expression of emotion in singing, and the relation between them, are interpreted, compared and analysed through an examination of the studies of authors whose research has been connected with the detection and analysis of the emotion in everyday life and in music performance. Comparing and combining visual and acoustical expressive elements of the performance of singing, the research in this thesis investigates the relative effects of one on the other and seeks to determine through empirical work which are crucial to the production and perception of emotional meaning in singing. These studies range from the conventional experiment in which the data are analysed statistically to individual subjective reports. The conventional experiments examine particular effects, whereas the subjective reports are used to address the more diverse properties of the performance.

The experiments range from descriptive to quantitative measurements of the expressive parameters of emotion where 'ecological validity' has been preserved by using realistic performance data. Two professional singers were filmed, videotaped and recorded in performance. Each singer performed in five different emotional conditions (anger, fear, happiness, sadness, and neutral). The performances (videotaped or real-time) were shown to fifteen audience members, who perceived facial and vocal expressions of emotion for each performer. Audience accuracy of the performed emotions was measured by comparing each performer's intentional expression with the audience's recognition of emotional meaning. Results showed a high rate of decoding accuracy of the performer's intended emotional (facial and vocal) expressiveness. A variety of empirical techniques were used including tracking and 'point-light' technique (Experiment-I and II), semantic differentials - Verbal and Non Verbal- (Experiment II and III), interviews with the performers and audience (Experiment-III and IV), and 'ecologically valid' real time performance assessment (Experiment IV). Experiment-I, by using point-light technique and digitised computer tracking and measurement, demonstrated that it is possible to differentiate emotional facial expressiveness in singing by using a purely quantitative-measurement technique (that is, without recurring to the subjectivity of the observer). Experiment II, by using point-light technique and semantic differentials (with descriptive emotional terms), demonstrated that kinematics alone provide enough information to distinguish between different expressive manners in the performer's facial behaviour when singing with different emotional meaning. Experiment III (videotaped performances with emotional content), by using semantic differentials and interviews with the performers, showed a high degree of consistency in the expressive elements (acoustical and visual) across repeated performances within the different emotional conditions. Experiment IV (real time performance), by also using semantic differentials and interviews with the performers and audience, showed a high degree of consistency in the expressive elements (acoustical and visual) across repeated performances within the different emotional conditions. It also revealed that musical structure and the performer's intention to enhance emotional meaning are important co-determinants of the communication process.

Chapter 1 presents the theoretical background on music/performance and emotion. Chapter 2 presents the theoretical-practical background on music/performance and emotion. Chapter 3 presents the empirical background on music/performance and emotion. A final aim of the studies has been to elaborate an expressive analytical tool that provides singers, singing teachers and students of singing with a reliable audience feedback about their capacities to communicate emotional meaning whilst singing. Since the author is both a performer and teacher it was essential for the investigation to integrate theory and practice. Therefore, a significant focus was also on the development of a recital, which aimed to stimulate debate for the theoretical and empirical results of this thesis. In fact, though performers have shown a partial knowledge of the expressive devices they used, and though all the experiments (videotaped or real-time) showed a high rate of the audience's decoding accuracy of the performance's intended emotional expressions, Experiment IV revealed that the audience's recognition of the performed emotional meaning would increase significantly if the performer, by having access to the audience's cognitive feedback, was able to check and improve the accuracy and consistency of the expressive cues used in the performance. All this experimental research is explored in Chapter 4. Chapter 5, the final chapter, presents a summary of all the empirical results. The thesis concludes with a brief discussion of further possibilities for research in the area and with the practical and theoretical implications to be drawn from this investigation.

TABLE OF CONTENTS

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

List of Figures

LIST OF APPENDICES

LIST OF CDs (DVD)

CD 1 – Shows four (4) still images of several framed moments of the point-light displayed movements illustrating different emotional conditions corresponding to Toni's sung versions of the phrase 'Mein Vater, mein Vater' of Schubert's song 'Erlkönig': Dscf0049.jpg.Fear; Dscf0052.Happy; Dscf0053.jpg.Sad; Dscf0054.Anger.

CD 2 – Shows ten (10) videotaped excerpts of the sung phrase 'Mein Vater' from Schubert's song 'Erlkönig', in the 5 different emotional conditions (Happiness, Sadness, Neutral, Anger, and Fear) by two professional singers: Silvia and Toni.

Silvia: S1POC.MOV.Happy; S1POC.MOV.Anger; S1POC.MOV.Fear; S1POC.MOV.Neutral; S1POC.MOV.Sad;

Toni: S1POC.MOV.Happy; S1POC.MOV.Anger; S1POC.MOV.Fear; S1POC.MOV.Neutral; S1POC.MOV.Sad.

CD 3 (DVD) – Shows ten (10) trials of the different performances of the Schubert's song 'Die Post', sung by Toni.

CD 4 – **Includes** the software designed especially for Experiment III and IV

CD 5 – **Includes** the recorded performance of the songs of *The Second Book of Winterreise – F. Schubert*, presented at Sheffield Anglican Cathedral, 17th March 2003.

INTRODUCTION

INTRODUCTION

Focus of the thesis

This thesis addresses the issue of performer expressiveness, focusing on the interface between singing and emotion. Singing, as an artistic performance, involves many skills: musicianship, technical vocal production and dramatic imagination, languages, control of non-verbal communication and expressiveness. All these requirements make singing perhaps one of the most difficult performing arts and, unfortunately, the mastery of all these elements has often resulted in a lack of performance expression. According to Salmon & Meyer (1992), performer expressiveness seems to be a primary element in effective performer-audience communication. So, an objective evaluation of the role of emotion communication within the performance of singing is an important element to see how a singer connects with the audience.

A main interest of this research project, thus, was to investigate how singers communicate emotion to their audience, and to understand some of the connections between music and emotion that could make young singers become aware of the importance of the expression of emotion in singing. It is important that young singers become aware of the role of emotional expression whilst singing because, as suggested by Balk (1985), most of the problems experienced by singers while communicating expressively often occur because they remain mainly focused on musical notation and technical skill, lacking expressive knowledge. Though they know they are supposed to communicate emotions to an audience whilst singing, it seems that they do not know how to do it, and/or restricted in how they may approach the task. In this sense, one important aim of this thesis became to develop an "expressive tool" that could investigate the area of emotion communication in that it might help young singers to sing more expressively. As a performer and teacher it was essential for me to integrate theory and practice. Therefore, a significant focus is also on the recital, which aimed to stimulate debate for the theoretical and empirical results of this thesis. Whilst the main area of interest of this thesis remained concerned with the audience perception and feed-back of performance expressiveness, other topics like the importance of facial versus vocal cues, objective versus subjective measurements, and laboratory

versus ecologically valid experiments were also investigated. An evaluation of audience perception and recognition of the intended emotional meaning was an additional aim of the study.

GENERAL BACKGROUND

According to Reber (1985), the word emotion is derived from the Latin *emovere*, meaning to move, excite stir up, or agitate. A common usage of the word 'emotion' can be understood as a physical experience that can be different in *quality* (pleasant-unpleasant), *intensity* (high-to-low), and *valence* (positive-to-negative) and occurring at a certain short period of time. Nevertheless, it seems difficult to present a simple and concrete definition of emotion as there are as many definitions as there are theories (LeDoux, 1996; Damasio, 1994; Frjida, 2000; Panksepp, 1982; Plutchik & Kellerman, 1989; Russel, 1980; Izard, 1977; Scherer, 1993; Ekman, Friesen & Ellsworth, 1972; for a comprehensive review in emotion literature see, for example, Ekman & Davidson, 1994; Strongman, 1996; Lewis & Haviland-Jones, 1993).

However, and despite the differing theoretical orientations, there are some aspects of emotion, which are generally accepted:

i)   A stimulus, either external or internal, triggers an emotional response;

ii)   There is a physiological component to emotion which may or may not be involved;

iii) Emotion motivates behaviour, either by instigating action or regulating behaviour;

iv)   Emotion often involves cognitive appraisal either before or after the emotion occurs, but emotions may exist purely as subjective experience (Ekman & Davidson, 1994; Lewis & Haviland-Jones, 1993).

So, for a working definition, it seems possible to agree with Peretz (2001) that emotion may involve three different levels of elements: experiential, neural and expressive. The theorising of emotion is, however, not a major concern of this thesis. The emphasis is on the processes of emotional expression and recognition, and how these processes can be

developed and improved in order to communicate the intended emotional meaning within the performance of singing. Thus, the expressive element of emotion is highlighted. Reber (1985) has defined expression as an outward display, connected or not to felt emotions, but often implying them. Additionally, Ekman (1993) uses the term "facial expression", in which 'expression' is meant to be a key component of the emotion displayed. According to Holt (1996), an expression of emotion to be intentional requires a *medium* and a *message* and, also, some sort of *interaction* with others. The idea that an expression of a certain emotion can be intentional seems to suggest that an intended emotional meaning can be worked out and improved actually while performing and communicating to an audience. It also implies that there is a shared *code* of emotional communication between the performer and the audience.

The investigation of the expression and communication of emotional meaning in singing in this thesis remains confined to only four of the basic emotions: happiness, sadness, fear and anger. The reason for the choice of these emotions is the universality and innateness of their expression and recognition, which are hypothesised to be associated with distinctive innate response patterns and neural substrates (Ekman & Friesen, 1972, 1978; Izard, 1971, 1977, 1992; Ekman, 1992). According to Caccioppo et al. (1993:134), these four above referred emotions (plus disgust) "are generally considered the basic emotions, but fewer than these five emotions (Panksepp, 1982), and more than these five emotions (Fridja, 1986) have also been suggested." Disgust, however, has not been considered here because it does not seem to be part of the ensemble of basic emotions expressed in the core singing repertoire (Nineteenth Century German Lied). Surprise, is often considered also as a basic emotion (Frijda, 1994), but has also not been used in this investigation. The reason for this lies in the fact that surprise seems to appear mostly to blend some other sort of emotional content (fear, happiness). In fact, surprise can be considered as a 'proto-emotion', because as Sloboda explains "if surprise is appraised as pleasurable, then an emotion of joy or elation may emerge" (Sloboda, 2001:93). Whilst evidence points to the existence of basic emotions and corresponding innate and universal facial expressions (Ekman, 1989; Izard, 1980, 1994), Russel & Fernandez-Dols (1999) suggest the existence of a bi-directional feedback between innate and cultural, or social, learned emotional expressions, in such way that each may influence and slightly modify the other. Whatever the case, from the

performer's point of view, and in order to achieve the best effect in the communication process, the intended emotion has to be expressed in a *clear, recognisable* and *authentic* way so that the audience can identify the basic emotion expressed. *Clear* means that the basic emotion expressed by the performer should not unintentionally blend any other emotional expression, and should avoid every other non-verbal communication (voice tone, stress, bodily tension, intonation, facial gesture, body movement) that could unintentionally disturb the clarification of the emotion expressed. On the contrary, as suggested by Bull (1983), if the performer's bodily movements enhance the expressiveness of the musical structure, the audience will have an accurate and clearer perception of the performance. This suggestion, the relevance of Non Verbal communication involved in the recognition of musical meaning, is investigated in this thesis by considering the impact of facial expressiveness on the communication of music emotional meaning. By *recognisable* it is meant that the emotion should be expressed according to social and cultural code common to the performer and the audience. As for *authenticity*, it is possible to understand this in two complementary ways:

> - One authentic way, concerns the fact that the expression of emotion is displayed and easier identified when it is supported, by the corresponding felt, thus, genuine emotion. In fact, an important conclusion of the research in this thesis, corroborating Runeson & Frykholm's (1983) demonstration that covert mental dispositions are specified in movement and can be detected by observers, suggests that the audience can clearly detect the authenticity of the emotional communication. Thus, the more an intended emotion is expressively felt, the clearer it will be to the audience to perceive its expression as a component of the emotion displayed, and more easily it will be to understand its 'emotional meaning'. This seems to corroborate the idea that that there is a necessity among performing musicians to generate 'true' performances in order to communicate with the audience (Persson, 1995). Bach (1778), for instance, argues that "a musician cannot move others unless he too is moved." This need and capacity of performers to generate emotional meaning in music performance - "to generate effectively an emotionally based conceptualisation" - is what Persson (2001:286) names, *musical reality*.
>
> - The second authentic way, concerns the fact that, in the performance of music, the emotion expressed by the performer should underline and corroborate the

emotional meaning of the musical structure in order to achieve its greatest authenticity. Of course, the perception of emotional meaning in music is different to the perception of expressed emotion in live performance, because live performances constitute a triadic relationship between the music, the performer, and the audience. The performer assumes the role of representing the music, and of delivering it according to the demanding artistic expressive communication. Finally, the performer is responsible, as said, to interpret the music according to the emotional meaning structurally presented by the musical discourse itself, which ability to communicate emotion and to arouse feelings in the listeners, has been acknowledged since Plato (Plato, *The Republic*, Book III, in Jowet, 1952:399).

In this sense, the content of this thesis will also present a review of some of the key conceptual tools implied in the understanding and theorising of the relationship between music and emotion; a review of the theoretical-practical background for empirical approaches to the study of the expression of emotion in music/performance will also be included; and a review of the main empirical evidence on the expression of e-motion in music/performance.

Outline of the chapters

Chapter 1 reviews some of the key concepts for the theory and research of the relationship between music and emotion. The debate between formalism and expressionism, between absolutism versus referentialism, and the different nuances of the debate between absolute expression and referential expression. The key concepts of structure, meaning, expressiveness, symbol and sign are reviewed according to different theoretical contributions and in light of recent findings of the new performing arts research field. The *iconic* contribution of the philosopher Langer (1942); the *autonomic* contribution of the formalists Hanslick (1854), Pratt (1931), Scruton (1997) and of the absolute expressionism of Meyer (1956); the *structuralist* contribution of Schenker (1935) and Ruwet (1967); the *symbolic* contribution of Frances (1958), Nattiez (1990), and Shepherd & Wicke (1997); the *phenomenological* contribution of Clifton (1983); and the *ecological* contribution of Gibson (1979), Bregman (1990) and others.

In **Chapter 2**, an historical perspective is presented exploring the rapport between music (singing) and emotion, and a review is made of some of the most important frameworks developed for empirical approaches in the study of the expression of emotion in music performance (singing): the 'Art of Content' model and 'Art of Form' model according to the way emotion may be conveyed in music/performance (Noy, 1993); the 'Semiological' model (Nattiez, 1990; Shepherd & Wicke, 1997; Dowling & Harwood, 1986) and the 'Listener-centric' (Lavy, 2001) model according to the way an emotional response to music/performance is evoked; the 'Functionalist' model (Juslin, 1995, 2001) according to the way emotion is communicated in music/performance; and the 'Ecological' model according to the way e-motion is perceived in music/performance (Shove & Repp, 1995). Of course, most of these models overlap in their approaches. Nevertheless, it is the ensemble of the different perspectives and the debates around them that supposedly cast light on the understanding of the different relationships between music/performance and emotion.

In **Chapter 3**, after a short review of the empirical evidence on music/performance and emotion, a semiotic review of the 'expressive musical event' is presented, where the empirical evidence on the expression of e-motion in music/performance is classified according to several semiotic levels: the *Indexical-qualitative coding*, through which the emotional character of music (the expressive event) might be partly connected to the perception of the articulatory and interactive movements of the performer, and partly connected to the acoustical structure of small segments of sounds (timbres, for instance); *the Iconic-dynamic-introversive coding*, through which the expressive event might imply specific and conventionalised structural and transformational features of the performed music 'symptomatic' of human emotional states; the *Iconic-dynamic-extroversive coding*, through which the expressive event can be directly connected to a 'virtual' source of emotion emergent from the 'trompe l'oreille' of an intended group of sounds dependent on the sound-structuring sequence generated by the performer's articulatory movements; and, the *Symbolic-associative coding*, through which the expressive event might remain indirectly connected to the articulatory movements of the performer, and to all the factors of the music that can be connected with emotional content in an individual's or collective's

memory, often due to occasionally and idiosyncratically learned associations and conditioning.

Chapter 4 reports the empirical studies undertaken. **Experiment I** focused on the study of the emotion facial expression in the performance of singing in a quantitative-objective way. To do this, the singer's facial gestures and how they were connected to the expression of emotional meaning in singing were investigated. This was done using a purely quantitative approach (computerised tracking and measurement of facial movements during the singer's different expressions of emotional meaning) without reference to the subjectivism of the audience's perception. **Experiment II** focused on the study of the facial expression in the performance of singing in an objective way, through the perception and recognition of point-light displays of the different emotional expressions performed. Thus, how and if a certain audience perceives and recognises the emotional meaning of the dynamic information contained in the abstract movement of the facial muscles recorded as reflective markers (point-lights) on film. **Experiment III**, focused on the study of the perception and recognition of emotional meaning in singing, through the observation of videotaped performances of singers singing the same musical phrase with different emotional intentions. Thus, an empirical and objective measure was obtained of how a certain audience perceives and recognises the emotions communicated through the voice and the face of the singers in different videotaped performances of singing. **Experiment IV** focused on the study of the perception and recognition of emotional meaning within the performance of singing in a real concert situation. That is, how a certain audience perceives and recognises the emotional meaning communicated through the voice and face of the singer in the performance of singing.

In **Chapter 5**, final conclusions on the empirical evidence on singing and emotion are presented, together with theoretical and practical implications, and the limitations of the present investigation, as well as a set of ideas for subsequent studies.

The songs from *Winterreise* also constitute the performance element of this thesis, as a practical application of the investigations here undertaken (CD 5 – includes the recorded and videotaped performance of *The Second Book of Winterreise - F. Schubert*, presented at

Sheffield Anglican Cathedral, 17th March 2003). The main focus of this study is thus connected to the facial and vocal expression in singers and the interpretative characteristics of the Lied repertoire, which oblige the singer to express the musical meaning almost exclusively through these two main channels of communication. This is the reason why the examples of the singing repertoire chosen for the empirical investigation have been all chosen from the Lied repertoire. This is also the reason why the elaboration of a 'cognitive feedback tool' in music/performance (singing) to assist singers and singing teachers to check and improve the expressiveness of their performances is one of the practical aims of this thesis. In this sense, Experiment-IV based on the performance of Schubert's Lied 'Die Post' served as a pilot-study for the interpretation of the *Winterreise* songs in the PhD final recital. Therefore, the Lied was not only previously analysed in itself, but also as an integrated part of the general context of the song cycle (see Appendix F). So, the hypothesis underpinning of Experiment IV was: if, during the course of this study, the perceivers' recognition of the singer's emotional interpretation could enable the performer to have a clearer perception of his/her emotional expression and, consequently, to help him/her to improve the expressiveness of his/her performance, then it might be possible to conclude that the process used in this experiment would have accomplished its aim and that it might aid other singers and singing students to improve the way they use their voices and their faces when interpreting emotionally the structurally content of a musical piece. And, consequently, that it could be used to aid the study and interpretation of the *Winterreise* songs presented in the recital.

CHAPTER 1

1. THEORETICAL BACKGROUND

REVIEW OF SOME KEY CONCEPTS INVOLVED IN THE STUDY OF MUSIC/PERFORMANCE (SINGING): 'EXPRESSION', 'MEANING', 'INTENTION', 'STRUCTURE', 'REFERENCE', AND 'PERCEPTION'.

1.1 INTRODUCTION

This thesis focuses on the Western Art Music Canon and is primarily concerned with solo singing (within the Nineteenth Century German Lied repertoire) involving verbal and/or dramatic action. Singing is constituted of an audible voiced melody born out of an embodied and expressive practice, the musical performance, within a process of meaning communication. Through the encounter of the voice and the music on a bodily emanation of meaning, singing becomes the core of a communication process, of a dynamical interaction between the individual and the world. Though always considered within its musical context, the emotionally expressive character of singing with the words and/or a dramatic context, usually with an explicit narrative, is taken together and considered as a complex yet integrated whole.

The most reasoned way to define music and its performance is to highlight the multifaceted nature of the musical experience and the plurality and relativism of its expressed meaning(s). As Clarke (1995) points out:

> "To finish, I return to the multiplicity of functions and diversity of significations that characterise expression in performance, at the same time acknowledging that this contribution has itself fallen into the trap of focusing too exclusively on the relationship between structure and expression. A host of equally significant perspectives warrants further investigation: the relationship between music and motion and the involvement of the human body (see Lidov, 1987, and Shove & Repp, 1995); the social dynamics of performance, both between co-performers and between performers and audience; the visual component of expression in live performance, a field open up by Davidson (1991, 1993); and much else. No doubt we will continue to obtain a variety of necessarily partial views of this rich and intriguing manifestation of human creativity, but, with increasing recognition of the possibilities of acquiring a more multidimensional perspective, perhaps progress can gradually be made towards a less fractured understanding."
>
> (Clarke, 1995:53)

The 'multiplicity of functions' and the 'diversity of signification(s)' are well culturally accepted and recognised attributes of music and its performance. Still, one should not yield

to the temptation of considering music, in and of itself, as a 'degree zero' (Barthes, 1989) of meaning and communication, as an 'empty form' (Eco, 1972), or as an 'empty sign' (Kristeva, 1989) ready to be filled up by any 'opportunistic' meaning.[1] In this sense, music should not be conceived as an object living in a timeless glass-case, away from its psychological, social and cultural constitution, and away from its impact on reality. As Shepherd (1997:8) points out: "to think music as a cultural process is to imply that social or cultural elements are contained within or passed through its sonic components." A long tradition of music aesthetic and musical meaning theory, like formalism, has avoided the relationship between musical processes and processes of subjectivity, claiming instead an autonomous nature and value of music, thus creating an irreconcilable dualism between intra-musical and extra-musical elements. As Kerman (1985:72) argues "...by removing the bare score from its context in order to examine it as an autonomous organism, the analyst removes that organism from the ecology that sustains it." Music, therefore, should be conceived generally, as a signifying practice, a process of communication among people and societies, a powerful means of meaning production and knowledge, subjectively, socially and culturally constituted. Thus, within the context of this thesis, singing and its performance practice will be considered as a particular instance of music as a signifying practice.

The most common shared idea amongst the different research approaches to expressiveness in music performance considers the concept of 'musical expression' as being related to and presupposing the concept of 'musical meaning'. Nevertheless, it seems that no stable consensus has yet been achieved, among theorists, about what constitutes 'meaning' and 'expression' in music. This lack of consensus can be found, first and foremost, in the metaphysical (philosophical and aesthetic) legacy that both concepts still carry with them. Musical meaning has been subjected to the most controversial debates. The principal and persistent arguments can be traced from the more contemporary discussions between

---

[1] According to Barthes (1989:299), music "constitutes a kind of *primal state* of pleasure: it produces a pleasure one always tries to recapture but never to explain; hence, it is the site of a pure effect... an effect severed from and somehow purified of any explicative reason, i.e., ultimately, of any *responsible* reason." Eco (1965), also, considers music as an 'empty form' to which one can attribute many possible meanings (see Nattiez, 1990:21), while for Kristeva (1989:309) music "is a system of differences that is not a system that means something, as it is the case with most of the structures of verbal language, (...) it is an "empty sign".

'absolutism vs referentialism' or 'formalism vs expressionism', back to the 'enlightened' - 'rationalism vs empiricism' - quarrel. As Bowman (1998) comments:

> "Enlightenment rationalists (…) shared the conviction that beneath music's kaleidoscopic surfaces and seductive charms lay some kind of orderly and rational principle accessible to human logic and reason. (…) To the rationalist, musical experience could not be a matter of blind, sensual pleasure and indulgence. Rather, music is orderly, patterned, systematic, the product of rules and principles. (…) Where rationalists regarded music as a sensory intellectual experience, empiricists considered it mere sense. To the ardent empiricist, musical experience was a response to a stimulus more psychological than logical. Musical responses (…) are mediated by faculties of 'imagination' or 'taste' that are not so much functions of rationality as of sensory pleasure. Thus, empiricist accounts of music emphasised its sensory side, its capacity to arouse emotion: musical pleasures were more bodily than mindful. (…) To the empiricist, music was not rational, but emotional and sensory - a matter of sentiment not of judgement."
>
> (Bowman, 98:72-74)

Following on from the Empiricists vs Rationalists quarrel, the subsequent theoretical accounts of musical meaning kept enhancing the dualism of its categories and their mutual exclusivity. Mind and body, subjectivity and objectivity, inner life and outer world, reason and emotion, intellect and feeling, universal and particular, culture and nature became the binary opposites and the conceptual evidence of the dualistic approach of Idealism. The dualistic approach of these categories, some of them can be traced back to Plato (Plato, *The Republic*, Book VII, in Jowet, 1952:531), still contaminates the present theoretical thinking and the conceptual tools that make part of the most up-to-date research on musical meaning. In this sense, a critical review of the main theoretical approaches to musical meaning will check now the particularity of the categories used in order to clarify and enlighten the conceptual tools that lead the contemporary research concerned with the meaning of music/performance.

1.2   THE ICONIC CONTRIBUTION

The dualistic conceptions discussed above can clearly be found in direct followers of 'idealism', like Langer (1942), who claimed that music is concerned not so much with 'feelings-felt' as with the 'form of feeling'. Like many idealistic thinkers, Langer treats musical meaning as 'universal form'. According to her, "music-heard and feeling-felt both

have formal characteristics and qualities that can be traced to a universal form" (Bowman, 1998:222). In fact, music for Langer does not really symbolise feelings, which occur, but the 'concept' of them; and, does not express emotions but 'conceptions' of emotions. In the last instance, music symbolises the 'form', the 'abstract form' common to all feelings. Davies' (1994) commentary on Langer's explanation of the expression of emotion in music makes us aware of the problems, dangers and difficulties of how to produce an accurate and scientifically valid explanation of musical meaning based on the metaphysical legacy of idealistic concepts:

> "Langer's theory removes emotion from art, replacing it with conceptions of emotions. In doing so, her theory undermines the basis for emotional responses to musical works and makes mysterious the power of art to evoke such responses. An account of musical expressiveness that fails to explain why we find music moving must be significantly flawed, and Langer's is such a theory."

> (Davies 1994:134)

The work of Langer is, nevertheless, a very sophisticated and elaborated one, and some of the concepts and ideas she has developed are still a fundamental part of the background knowledge to understand the interface between 'the inner life' of feelings and the 'outer world' of music experience. Her most important device, the concept of music as an 'unconsummated presentational symbol', is however not an unproblematic and uncontroversial one.[2] This 'unconsummated presentational symbol' constitutes an important conceptual device in her attempt to explain how music can convey emotional meaning without actually denoting, representing or expressing any determinate feeling beyond it. Through it, Langer has created a new kind of symbolic situation, i.e., a "purely connotational semantic" (Langer, 1942:101). According to her, it is understandable to say that music expresses feelings only because music's form and feeling's form *present* an ideal structural similarity. As Davies (1994:133) puts it, "in appreciating musical works as expressive, we are appreciating them as presentational symbols; that is, as conveying

---

[2]   It is not the intention here to get involved in an exhaustive and further detailed examination of the philosophical explanations and controversial theories of this author or other subsequent quoted authors whose work is also concerned with emotion and musical meaning. The intention, here, is, in light of recent research and evidence, to recover and to refresh some of the ideas and concepts created and used by these authors, and to explore them as a tool to help understand and conceptualise contemporary findings and perspectives. An extensive bibliography whose aim is a detailed philosophical discussion on this subject can be found in authors like Budd (1985), Davies (1994), Walton (1970), Bowman (1998), Scruton (1997), Robinson(1997), Small (1998), Kivy (1989).

conceptions of emotions". Thus, what people perceive as feelings and emotions when they hear music is "a tonal analogue of emotive life" (Langer, 1953:27). Langer's explanation is that, what people tend to perceive in music as moods and emotions are just abstract forms of emotions that, according to Pratt's (1931:230) most famous slogan, 'merely *sound* the way emotions *feel*'.[3] According to Langer's theory, music *presents* emotions and feelings, but does not *refer* to them. As Bowman comments (1998:220), Langer's presentational symbol stresses the fact that the emotive significance of music when understood as such a symbol "is not a function of feelings felt, evoked, or aroused but of insight into the form of feeling." Music appears, thus, as an "unconsummated symbol" (Langer, 1942:241), it does not refer to particular feelings, and it has no assigned reference. On the one hand, Langer's rejection of music's symbolic reference leaves us only with the "ghost of meaning" (Howes, 1958:15). On the other hand, her claim that music has 'vital import' introduces an essential possibility of iconic mapping between the aural elements of a musical work and the dynamic elements of the inner life. Langer's own words are:

> "There are certain aspects of the so-called 'inner life' which have formal properties similar to those of music - patterns of motion and rest, of tension and release, of agreement and disagreement, preparation, fulfilment, excitation, sudden change, etc."
>
> (Langer, 1942: 228)

Commenting on this very same Langer's paragraph, Cook and Dibben (2001: 58) have two different remarks:

- One, concerns the fact that, "since Hanslick's day the idea that music expresses emotions through somehow mirroring the dynamics of our inner life has become a commonplace in writing on the topic."

---

[3] Though the psychologist Carroll C. Pratt and the philosopher Susanne Langer approved each other's work, it seems that Pratt did not share Langer's theory of music's 'unconsummated symbolism'. According to Davies (1994:134), Pratt does not hold the view that music is a (non-referential) symbol but that, it "possesses an objective, emotional character due to the formal similarity between musical motion and the organic or kinaesthetic sensations that are the correlates and determinants of emotions". "Music at its best is not symbolic at all...tonal design stands for nothing beyond itself. It does not suggest mood and feeling. It *is* mood and feeling. The qualities of auditory perception are not iconic signs, nor do they by themselves represent or copy anything' (Pratt 1954: 290, quoted in Davies 1994: 134).

- The other, related to the vagueness existing in such thinking, concerns the sort of entity that the 'dynamics of our inner life' might actually be, or how it can be measured.

Despite its vagueness, the idea of a 'similarity', or a 'homology', between the 'dynamics of music' and the 'dynamic qualities of feeling' has become an important topic of theorists and researchers all over the world. And, within Langer's theory, it is perhaps one of the most appealing conceptual tools that she has offered modern research towards understanding and explaining the expression of emotion in music. Indeed, the concept of 'vitality affects' presented by Stern (1985) may be considered as a reprise and direct application of Langer's conceptual tool. As Sloboda (2001) summarises:

> "Stern (1985) introduced the concept of vitality affects to describe a set of elusive qualities related to intensity, shape, contour, and movement. These characteristics are best described in dynamic terms such as crescendo, fleeting, explosive, diminuendo, etc. These qualities are not emotions, but rather abstract 'forms' of feeling that occur both together with, and in the absence of, proper emotions. The vitality affects are 'a-modal' in the sense that they are common to all modes of expression. Stern (1985) suggests that the vitality affects are of a particular importance in the early communicative acts of mother and infant. Mother and infant respond to one another by constantly adapting and adjusting the intensity, timing, and contour of their expressive acts. This process of constant matching of gestural events is referred to as *attunement*."

(Sloboda, 2001: 79)

Sloboda (2001:81) considers that "although the notion of vitality affects is admittedly vague, it seems to capture something important about music's expressiveness. (...) Vitality affects with their dynamic changes and patterns of events, are somewhat reminiscent of philosopher Susanne Langer's (1951) notion of music as expressive of the forms of feeling." As discussed above, Langer (1953:395) conceives music as presenting not particular emotions, not discrete feelings, but "articulations of feeling, forms of life and feeling, activity, suffering, selfhood", and also "patterns of motion and rest, of tension and release, of agreement and disagreement, preparation, fulfilment, excitation, sudden change" (Langer, 1942:228), in brief, the dynamic form of emotional life.

It seems, according to Davies (1994), that some ambiguity could be found in her theory in what is, effectively, symbolised by musical form. Sometimes she claims that what is symbolised is the form common to all feelings, i.e., not only different feelings but also contrary feelings. In fact, Langer (1951:202) claims that "some musical forms seem to bear a sad and a happy interpretation equally well". At other times, she affirms that what is symbolised is the "form common to the various instances of a particular feeling" (see Davies, 1994:127). Davies suspects that this ambiguity appears "because she realises that different feelings, even contrary feelings, might share the same form, while she also believes that some feelings have their own, distinctive forms" (Davies, 1994:127).

The first half of Langer's ambiguity might be explained if we understand the concept of feeling through the concept of 'vitality affects' introduced by Stern (1985). Thus, in a way, what music articulates as emotions or moods might be precognitive, 'a-modal' affects, i.e., which are common to all modes of expression. As Sloboda (2001:81) explains, the possibility of music to symbolise the form of different or, even, contrary feelings "might be because vitality affects are not emotions themselves, and because some vitality affects are common to many different emotions."[4] The second half of the ambiguity found in Langer's symbolism by Davies, might be explained through the 'prototype approach to emotions' based on the work by Rosch (1978). The 'prototype' is, as Sloboda (2001:78) explains, "an abstract image that consists of a set of weighted features that represent the exemplar of a family of instances." Indeed, some emotions are more prototypical than others are. For example, infatuation, or adoration is a better exemplar of love than caring.[5]

Whatever the explanation might be for the existence of this ambiguity in Langer's theory, the compelling reason, though, for this sense of vagueness and imprecision in such an appealing and important conceptual tool, like 'iconicity', to explain and connect musical meaning to musical experience, lies, again, in the idealistic ascendancy of her own

---

[4] From a very different perspective, a hermeneutic one, Henry Orlov (1981:137) explains that the reason for this, is that "the reality so symbolised is that of pre-verbal experience."

[5] For a further explanation of the Rosch's (1978) 'prototype approach', see Chapter 4, under the item 4.2 'Methodological Questions'.

definition of music as an 'unconsummated presentational symbol'.[6] Thus, Langer creates the idealistic concept of 'unconsummated presentational symbol' in order to be able to deny and to reject any attempt of classification of her theory as arousalist. According to her, emotion cannot literally exist in music. One of its features (its form) will have to make the connection to the universe of human feelings. The relation is, thus, just a symbolic but not a referential one; emotion is naturally possessed by music but not aroused by it. Therefore, it might be concluded that music expresses not emotions but, instead and only, conceptions of emotions.

1.3   THE AUTONOMIC CONTRIBUTION

Considered as "the other side of Idealism's dualistic currency" (Bowman, 1998:133), Formalism, treats musical meaning, nature and value as strictly intrinsic to musical work itself. According to the formalistic theory, musical meaning is inherent to music's internal structures. The Nineteenth Century defender of music's purely musical value, Hanslick, believed that "music is, first and foremost, objective structure" (Hanslick, 1957:32). These structures are materially occurring forms and shapes of sounds, which appeal to one's senses. They constitute what Hanslick has called 'musical beauty', which is, according to him, 'concrete, immediate, sensuous, and alive' (Bowman, 1998:146). Music is, thus, according to formalism, objective and autonomous, instead of arbitrarily subjective and heteronomous as 'referentialists' theories considered it. According to Hanslick's (1957) own words:

> "The forms which construct themselves out of tones are not empty but filled; they are not mere
> contours of a vacuum but mind giving shape to itself from within, and not depending on its
> effects on the audience, whether it is able or not to arouse emotionally the listener."
>
> (Hanslick's, 1957:30)

---

[6] According to Davies (1994:123-124), Langer considers music as "an iconic symbol of mental states identifiable as emotions. She regards meaning of music as relying importantly on a natural element, iconicity, and she denies that, as a symbol, a musical work depends for its significance on its place within a general symbol system." It seems, also, to be this the mind of Bowman (1998:210) who confirms that "consistent with the feelings of idealism, this (Langer's) explanation imputes a kind of controlled rationality to the way feeling is encountered in music. Thus, music can be a felt phenomenon, yet since it is not concerned with actual, particular feelings but with feeling in general, it is not a dangerous or sensually affair." Davies (1994:134) also sees the flaw in Langer's account of musical expressiveness owing to a dependence on her idealistic concern with rejecting the arousal theory of music expressiveness. The importance of the iconicity concept to explain and articulate musical meaning and music experience will be discussed further in item 1.5 "The Symbolic Contribution."

Though not denying the possibility of the existence of a referential meaning in music, the formalistic perspective stresses the importance of what is purely and strictly musical. Another prominent formalist, the psychologist Pratt (1954) presents an important twist to Hanslick's theory. Pratt argues that Hanslick had a too restricted idea of what a form may convey. Like Hanslick, he considers that music cannot contain or refer to emotions because these are part of the bodily process of the subject:

> "The material of emotion is a bodily process, an elaborate pattern of muscular and visceral disturbance."

> (Pratt, 1954:291)

Nevertheless, and as Davies (1994:135) points out, "because of its dynamic character, music presents formal properties with an objective character so similar to that of the emotions that it can be said to be able to possess or present an emotional character." In this way, music presents forms that become acoustically perceived correlates of organic or kinaesthetic sensations: agitation, restlessness, calmness, vacillation. As Pratt (1954) claims:

> "The emotions and strivings of the will and desire are embodied in music not directly, but indirectly by way of tonal designs which closely resemble in formal outline the inner movements of the spirit."

> Pratt (1954: 290)

Though Pratt seems to be willing to affirm that music is able through its character to create a kind of strong field of resonance with the emotions, his explanations of how music accomplishes that fact are unable to overcome the difficulties created through his dualistic approach to musical meaning.

Meyer's (1956) theory of 'absolute expressionism' in music might be seen as the first objective attempt to explain musical meaning from a formalistic perspective.[7] His theory of musical meaning is based on Dewey's (1894) conflict theory of emotions and it is also

---

[7] According to Meyer (1994:4), "an absolute expressionist is a listener who believes that musical meaning is intrinsic in the music, and who derives an emotional meaning from understanding the interplay of the music relationships."

connected with two other main psychological theories: Gestalt psychology and Information theory. According to Meyer (1956), two main different theories, 'absolutism' and 'referentialism' dispute an interpretation of the meaning of music. The former theory claims the intrinsic absolute meaning of music and considers that musical meaning "lies exclusively within the closed context of musical work itself", i.e., that musical meaning derives exclusively from the abstract relationships of music in and of itself. The latter claims the 'extrinsic' referential meaning and considers that musical meaning "refers to the extra-musical world of concepts, actions, emotional states and character" (Meyer, 1994:3).

Despite the fact that these two theories seem to differ radically in their musical content, it seems acceptable that absolutist meanings and referential meanings are logically compatible. Meyer's own perspective is primarily concerned with the meaning of the different musical processes that are "to be found in the structural interplay of music elements and forms" (Meyer, 1994:1). Nevertheless, Meyer sustains the opinion that these two different musical meanings are not only logically compatible but that they may be coexistent in one and the same musical piece:

> "In spite of the persistent wrangling of these two groups ('absolutists' vs 'referentialists'), it seems obvious that absolute meanings and referential meanings are not mutually exclusive: that they can and do coexist in one and the same piece of music, just as they do in a poem and a painting. In short, the arguments are the result of a tendency toward musical monism rather than a product of any logical opposition between types of meaning."
>
> (Meyer, 1994:4)

Thus, and again, despite the emphasis placed on the intrinsic meaning of music, Meyer accepts that referential meanings do exist in reality, and even (like purely musical meanings) they are cultural and socially learned. He writes, for instance, that "the musical theory and practice of many cultures in many different epochs indicates that music can and does convey referential meaning" (Meyer, 1994:6).

Meanwhile, the debate concerning musical meaning from the two opposing aesthetic positions of formalism and expressionism, has gained with Meyer's theory, a new ingenious insight: "thinking and feeling need not to be viewed as polar opposites but as different manifestations of a single psychological process" (Meyer, 1956:34). Formalism

has often confused expressionism and referentialism by believing that any emotional response to music can not happen without reference to some kind of extra-musical meaning. Meyer's approach to this issue considered the possibility of a response to music capable of involving emotional arousal by "meanings which lie within the closed context of the musical work itself" and without reference to the "extra-musical world of concepts, actions and human emotional states" (Meyer, 1994:6).[8] The explanation for this lies in the fact that:

> "Though (intellectual and affective responses to music) are psychologically differentiated as responses, both depend upon the same perceptive processes, the same stylistic habits, the same modes of mental organisation; and the same musical processes give rise to and shape both types of experience."
>
> (Meyer, 1994:34)

Thus, felt and intellectual responses to music are not different processes, but, as he concludes, "different ways of experiencing the same process" (Meyer, 1994:34). Based on Dewey's (1894) 'conflict theory of emotions', Meyer conceived that the arousal of affect, or emotion, happens "when an expectation - a tendency to respond - activated by the musical stimulus situation, is temporarily or permanently blocked" (Meyer, 1994:28). Thus, any deviation from the expected musical stylistic progression of sounds can be regarded as an emotional stimulus.[9] This will happen, however, only "if intellectual activity is allowed to remain unconscious. In this case, the mental tensions and the deliberations involved when a tendency is inhibited are experienced as feeling or affect rather than as conscious cognition" (Meyer, 1994:28). This leads to Meyer's concept of 'embodied musical meaning'. Meyer's 'embodied musical meaning' is not, like 'designative meaning', the product of the relationship between the musical stimulus and the events or consequences it indicates and which are different from itself in kind. In terms of 'embodied musical

---

[8] According to Meyer (1994:7) "the distinction just drawn between absolute and referential meanings is not the same as the distinction between the aesthetic positions which are commonly called 'formalist' and 'expressionist'. (...) One might, in other words, divide the expressionists into two groups: absolute expressionists and referential expressionists. The former group believe that expressive meanings arise in response to music and that these exist without reference to the extramusical world of concepts, actions, and human emotional states, while the latter group would assert that emotional expression is dependent upon understanding of the referential content of music."

[9] One should not misunderstand Meyer's arousal-plus-context formula. According to Cook & Dibben (2001:59), "the link he makes between psychological principles and musical meaning is not a simple or direct one. Meyer clearly recognises the extent to which listener's expectations reflect not only the properties of the musical stimulus but also the culturally specific stylistic codes that are acquired through enculturation."

meaning', the antecedent stimulus and the consequent event to which this stimulus points are of the same kind, i.e., musical.[10] But, as Meyer suggests, musical meaning is not to be found in either the stimulus, or what it indicates, or even, the perceiver:

> "Though the perception of a relationship can only arise as the result of some individual's mental behaviour, the relationship itself is not to be located in the mind of the perceiver. The meanings observed are not subjective. Thus the relationships existing between the tones themselves or those existing between the tones and the things they designate or connote, though a product of cultural experience, are real connections existing objectively in culture."
>
> (Meyer, 1994:30)

These two different accounts of meaning (designative and embodied), though recognisably coexistent in music, do not share, according to Meyer, the same importance and value, because extra-musical meanings (designative) are not as directly implicated by musical patterns as intra-musical meanings (embodied). Thus, Meyer's concerns about music's expressiveness are primarily directed to the way listeners, emotionally and/or rationally, make sense of the interplay of the musical structures and of the syntactical relationships among musical events. These concerns brought the idea of music expressing emotion within the objective relationships and structures of music, and created the opportunity of bringing emotion objectively into the realm of music and of music theory. Nevertheless, the strictly formal and cognitive parameters allowed by Meyer's absolutism to explain and research the expressive power of music and of musical meaning create more obstacles and difficulties then answers and solutions to this issue. As, Shepherd & Wicke (1997) explain:

> "To Meyer must be credited the 'humanization' of theory and analysis. He offers a bridge from music theory to historical musicology in the pursuit of criticism by putting people in theory's thinking. Yet, like most music theorists, Meyer's point of departure is not the dynamics of musical experience itself, but categories and procedures involved in music's creation and construction. Because the production of music is not unrelated to the experience it instigates, Meyer is able, convincingly and perceptively, to draw striking parallels between musical

---

[10] 'Embodied musical meaning' in the context of Meyer's theory implies that if one musical event has meaning, "it has meaning because it points to and makes us expect another musical event" (Meyer, 1994:31). Thus, the antecedent stimulus and the consequent event are alike in kind. Meyer's conception of musical meaning is based on Cohen's (1944:47) general definition of meaning ("anything requires meaning if it is connected with, or indicates, or refers to, something beyond itself, so that its full nature points to and is revealed in that connection"), and on Mead's (1934) triadic relationship: stimulus, consequent and observer (see Meyer, 1994:30).

processes and 'mental' processes. However, because such work is highly cognitive in character and likely as a consequence to be thin on protocols for connecting the sounds of music to music as an affective experience, Meyer is unable to make connections between the sounds of music and people sufficient to explain rather than describe musical experience. That is why Meyer experiences problems with the analysis of ethetic relationships."

Shepherd & Wicke (1997:143)

In fact, it seems that despite all Meyer's attempts and efforts to link the 'ethos' or 'character' of the music, it's 'feeling-tone', to formal and processive relationships and to ground ethetic responses in music's objective patterns, there is always something about ethetic relationships that *slips out* of the listener's mind. According to Meyer (1973:242), ethos is "quite literally felt, and it can be felt without the mediation of extramusical concepts or images." Such a kinesthetic sensing of ethos can only be explained through the listener's 'empathetic identification' with the music. A competent listener "perceives and responds to music with his total being", writes Meyer (1973:242). However, according to Shepherd's (1997:89) comments, "this ethetic relationship, which stands at the heart of musical apprehension for Meyer, poses considerable difficulty for analysis."

In Meyer's (1973) own words:

"...ethetic relationships are unquestionably important...(but) are hard to analyse with rigour and precision...[There is an] absence of an adequate theory of ethetic change and transformation."

Meyer (1973:245-6)

And again:

"...the analysis must end here...[because] the rigorous analysis of ethetic relationships is beyond my knowledge and skill."

Meyer (1973:267)

So, Meyer's work seems to oscillate between the will of having musical meaning under the structuring and powerful control of the mind, and the temptation to understand and explain why one's responses to the expressive power of the music, the 'music's character', implicate such a striking relationship to the body. Thus, the body and the 'ethos' remain irreducible to the categories and procedures involved in music's creation, construction and analysis preconceived by Meyer's theory, because they are not suited for the desired

explanation. Meyer theory's basic interpretative framework focuses primarily upon form and syntax. As Bowman (1998:179) points out, this makes him relegate to secondary status "musical parameters that defy quantification and notation, and experience that is dynamic, corporeal, and above all, qualitative."

All those elements that do not serve syntactical ends are considered by Meyer (1967:36) as merely 'sensuous' and therefore are 'evocative of sensuous-associative responses'. Music, according to Meyer (1967) must be evaluated syntactically because:

> "The syntactical response is more valuable than those responses in which the ego is dissolved,
>
> losing its identity in voluptuous sensation or the reverie of daydreams."
>
> Meyer (1967:35-36)

Like many notorious idealists before him, Meyer apparently believes, as Bowman (1998:180) remarks, "that musical experience consists in two mutually exclusive alternatives: syntactical understanding or mindless sensuousness." Thus, and despite Meyer's (1994:34) statements claiming to be adopting a position against the "traditional dichotomy between reason and emotion and the parent polarity between mind and body", it seems that the heavy metaphysical legacy of idealism is still at work within Meyer's theory of meaning in music, too.[11]

One last issue is still to be considered within Meyer's account of 'expressive emotional meaning'. Meyer (1994:7) claims that while referential expressionism asserts that "emotional expression is dependent upon an understanding of the referential content of music", absolute expressionism, as he advocates, believes that "expressive emotional meanings arise in response to music (...) without reference to the extramusical world of concepts, actions, and human emotional states." Based on Dewey's (1894) 'conflict theory' of emotions, Meyer attributes the emotional meaning expressed in music to inhibited response tendencies. However, critics claim that "conflict is neither necessary nor sufficient

---

[11] In this regard, Davies (1994:289) comments that "Meyer's views are vulnerable to many criticisms (...) he draws a sharp line between the emotional and the cognitive, regarding the two as exclusive." According to Budd (1985:158): "When a listener responds with emotion to the absolute meaning of a musical work, the stimulus situation is the music itself; and since the absolute meaning of music is intramusical, the emotional experience of the listener is intramusical. However, the conception of musical emotion put forward by Absolute Expressionism receives no support from Meyer's account of the nature of emotion – for this account has no foundation."

to emotion: there is emotion that is not caused by conflict, and conflict that does not generate emotion" (Bowman, 1998:176). Elliot (1987:26), for instance, relates that "affect is a continuous aspect of consciousness."

On the other hand, Meyer (1994:14) claims that what is "most vital and essential in emotional experience (is) the feeling-tone accompanying emotional experience, that is, the affect." Thus, by claiming that the behavioural and physiological constituents of emotion are mere adjuncts or concomitants of affect, Meyer reduces emotion to arousal creating the idea that affect is 'undifferentiated'. As Bowman (1998:175) explains:

> "The contention that 'affect' is undifferentiated is of pivotal importance to Meyer's theory, since it is the mechanism that purportedly allows for the structural equivalent of musical pattern and the feeling it disposes. Since affect in itself, 'emotion-felt', is undifferentiated (or more precisely, differentiated only in terms of intensity), it is possible, Meyer believes, to identify a range of 'expressive' musical experiences that are in no way referential or parasitic upon non-musical 'feelings'. Feeling (as opposed to particular, identifiable feelings) has to do with degree of arousal. Accordingly, whatever differentiation exists among felt musical experiences is created by objective patterns within the music itself. It follows that a profile of genuinely musical expression can be obtained from hard-headed analysis of the musical 'stimulus'."
>
> Bowman (1998:175)

It seems that Meyer's conception of emotion as an 'undifferentiated affect' relies mainly upon a quantitative criterion, ignoring almost entirely the role of the qualitative criteria in the evaluation, recognition and identification of expressed emotions.[12] It seems, though and contrarily to this, that a significant part of the present "psychological community shares the conviction that emotion is differentiated and multidimensional" (Bowman (1998:176).[13]

---

[12] This conception was based on Rapoport's psychological findings (1950:21). Meyer (1956:14) claimed that "emotion is undifferentiated" because "physiological adjustments cannot be shown to be sufficient causes for affective responses or, even, elucidate upon the relationship between affective responses and the stimuli which produce them".

[13] See Izard (1977), and Ekman & Davidson (1994) for a consistent conception of the emotion as differentiated and multidimensional. Nevertheless, the already referred Stern (1985) 'vitality affects' seem to offer an interesting contemporary perspective to the possibility of Meyer's undifferentiated affect. However, how relevant the case of 'vitality affects' be to a phenomenological description of music experience (see Imberty, 1997), these class of qualities seem not to have reached still a consensus of opinion about their status as emotions. On the other hand, Meyer's 'undifferentiated affect' is directly connected with the intramusical meaning conception, which does not explain, as in the case of Stern (1985) 'vitality affects' an extramusical, even if 'a-modal', emotional reference and awareness." For further discussion on the subject see Budd (1985:157-67).

As Hillman (1960:85) observes, "irritation and fury are not to be conceived as points on a scale reading from apathy to manic seizures, like degrees of temperature; they are rather individual conditions of the personality as a whole."

According to Cook (2001:58), the "relationship between emotions as normally understood - hope, joy, grief, and so on - and the kind of undifferentiated affect or feeling tone that Meyer's theory predicates" constitutes a serious difficulty of Meyer's approach. More generally, other theorists doubt whether Meyer's theory is appropriate to refer to the emotions at all, because there is a lack of consistency in the explanation of the associative process between intra and extra-musical meaning.[14] Davies (1994), however, has tried to offer an explanation of how for Meyer the possibility of the listener's differentiating the 'feeling tone' is a real one. According to Davies (1994), Meyer's position is possible by explaining that the feeling tone can itself take on a kind of 'mimetic' function:

> "There are, within any culture, groups of emotions having a common pattern of behavioural expression. For instance, the dynamics of the behavioural expression of sadness, grief, disappointment, and regret are similar and can be distinguished from the behavioural expression of joy, happiness and enthusiasm. Music can be heard as imitating the dynamics of behaviour. As listeners we hear these dynamics and imagine that the music is through them expressive of, say, sadness. In so doing we (imaginatively) make the music the emotional object in terms of which our feeling tone can be differentiated."

Davies (1994:288)

To conclude, one is left, again, with the main argument that *music expresses emotion through somehow mirroring the dynamics of our inner life.* Cook (2001), commenting on the same paragraph from Davies (1994) affirms that:

> "Davies is invoking the Hanslickean idea that feelings are characterised by specific dynamic qualities (...) and that music (...) by its motional qualities can specify adjectival qualities as the joyful or sorrowful quality of love. In effect, Davies is suggesting that Meyer's undifferentiated affect conveys such adjectival properties, which we are led to attribute to the music itself."

Cook (2001:58)

---

[14] At this regard, Budd (1985:172) also claims that "the basis of the association between the syntactical character of a musical work and the facts of human existence that Meyer highlights is not specified, and in consequence, it is impossible to determine whether the association is a well-founded and relevant to the music's value.

And, also:

> "The listener's feelings of impatience and frustration, while poorly mirroring the expressive
> structure of the work, may nonetheless provide the initial clues that something is
> amiss...providing an impetus to interpretation."

<div align="right">Cook (2001:61)</div>

But, it seems to be possible to take more than adjectival properties or interpretative clues out of Davies explanation. Implicit in Davies explanation is the claim that music not only involves the imagination as triggered by the music, but also that the imagination itself is directed and controlled by the musical context:

> "The listener's response not only is initiated by the music but also depends on attention to the
> course of the music for its elaboration and articulation. Where such a theory emphasises the role
> of imagination, that role is not of free association, but of thoughts (entertained without belief)
> prompted directly by the music and responsive to all its subtle nuances."

<div align="right">(Davies, 1994:291)</div>

Callen (1983), also claims that our experience of music involves imagination, which is directed and controlled by the musical context. Music becomes the emotional object of our emotion as a result of an act of the imagination. But, the work of the imagination is not to be seen without control or with free rein. The exercise of the imagination is, according to Callen, controlled purely by musical features and by the dynamic movement of music. As Davies (1994) explains:

> "There is a need for similarity between the dynamic pattern of the characteristic behavioural
> expression of the emotion we imagine into the music and the movement of the music. Music
> encourages the listener to move in a way that resonates with its own dynamic character. The
> ways we move affect our feelings (as well as vice versa); the movements we make, and their
> affective content, model for us the expressive character we hear (in a stylised form) in the
> music. When now the imagination contributes to the process, we have a sympathetic
> identification not solely of expressive quality but of (fictional) human feeling in the music."

<div align="right">(Davies, 1994:291)</div>

Different authors have explored the idea that music has not only a dynamic vitality similar to emotional life but also that the dynamic character of the music encourages to think of it as capable of expressing motion or movement.

Among these authors, Scruton (1997) deserves a special mention in the context of the autonomous accounts of musical meaning, specially because of the insight his theory brings to the relationship between the two concepts under discussion: 'expression' and 'meaning'. The statement that musical meaning is "essentially connected to an experience of movement", is a major issue in Scruton's (1997:55) theory. Statements like 'whenever we hear music, we hear movement' appear repeatedly along his arguments. His first move is to bring together the concepts of 'musical expression' and 'musical meaning' by placing the *sense of movement* at the centre of musical experience. Scruton (1997) sees the expressive character of music as a sympathetic response of the listener to the moving forms of the music, as:

> "A kind of latent dancing - a sublimated desire to 'move with' the music, and so to focus on its moving forms. (...) The experience of musical form is an experience of movements and gestures, detached from the material world, and carried through to their musical completion. In hearing the content of a piece of music, therefore, we are also hearing the form: the life, which grows and fulfils itself in tones."

> (Scruton 1997:357)

This account of musical meaning has a clear resonance with the Hanslickean description of the essence of music as "tönend-bewegte Formen", an obscure phrase that Scruton (1997:353) conveniently translates for "forms moved through tones". According to him, the formalism of Hanslick's most famous phrase is an acknowledgement of an 'irreducible metaphor':[15]

> "This might seem to confirm the judgement that Hanslick is a formalist, on account of the prominence given in the definition to form. But form is mentioned only in the context of movement. And this idea of musical movement is an irreducible metaphor, which can be explained only through our response to music. It is associated with other metaphors - and in

---

[15] See Scruton's (1997:80) explanation: "By metaphor I shall mean what Aristotle meant: the deliberate application of a term or phrase to something that is known not to exemplify it." And, Scruton (1997:92): "The metaphor cannot be eliminated from the description of music, because it defines the intentional object of the musical experience. Take the metaphor away and you cease to describe the experience of music. (...) The indispensable metaphor occurs, when the way the world seems, depends upon an imaginative involvement with it, rather than our cognitive goals. And this is the case when we listen to music." See, also, Scruton (1997:153): "Metaphor is the experience of 'fit' between two mental contents". For the important role of metaphor in human understanding, see also Lakoff and Johnson (1980).

particular with the metaphor of life. In hearing the movement in music we are hearing life - life conscious of itself; and if, sometimes, we use words like 'expression' to convey the character of this life, is this not merely a natural extension of the metaphor? In short, Hanslick has given us no alternative to the theory that he criticises: on the contrary, he has tacitly accepted its most important claim - that music is the object of a metaphorical perception, whereby it is lifted from the physical realm of sound and placed in the intentional theatre of our sympathies."

Scruton (1997:353)

Thus, Scruton's approach provides an explanation of musical meaning in which our response to music should be understood as a particular experience of movement in which he includes the perception of those movements characteristic of human emotion. Meaningful is the idea that the experience of movement that musical expression evokes to the listener's perception is, according to him, related to an 'acousmatic' or 'phenomenal space', founded in a metaphorical transfer and 'detached from the material world'. The existence of such a phenomenal space of music is based on the three important distinctions he places between sound and tone:

"that between the acoustical experience of sounds, and the musical experience of tones; that between the real causality of sounds and the virtual causality that generates tone from tone in the musical order; and that between the sequence of sounds and the movement of the tones that we hear in them. These distinctions are parts of the comprehensive distinction between sound and tone."

(Scruton, 1997:19)

These three distinctions make part of an important 'step' in Scruton's argument that musical meaning is what we experience through the music when we respond 'sympathetically' to the movement we hear in the sounds of the music. It is important in Scruton's argument that these motion experiences exist in a 'realm' quite different from the real world of the listener's everyday experience. The realm where the experiences of motion are supposed to be perceived is, according to Scruton (1997:75), the 'phenomenal space of tones', which is completely 'detached from the material world' and ontologically separated from the real spaces of sound:

"The apparent spatiality of the acousmatic realm is of a piece with its transcendence (with the fact that we encounter it as something other, something observed, rather than as an inner process). There is no real space of sounds, but there is a phenomenal space of tones."

(Scruton, 1997:75)

And, again:

"When we hear music, we do not hear sound only; we hear something in the sound [...]."

(Scruton, 1997:19)

Thus, it seems that, while denying the reality of the real things (the perceptual organisation of sounds in the auditory array of listener's everyday experiences), Scruton tries to proclaim the transcendence of their musical completion in the apparent spatiality of the 'acousmatic realm'. Though he affirms that it is not possible to advance from the phenomenal space to an objective spatial order, this apparent spatiality has, according to Scruton (1997:74), an objective character, once it is heard as objective:

"I do not encounter the Brahms Concert (Brahms's Second Piano Concerto in B flat major, Op. 83) as an inner process of mine, like a sensation or an image; it is out there, independent, and could be other than it seems to me."

(Scruton, 1997:74)

This is the reason why, according to Clarke (2000), we should consider Scruton's approach as aligned with the aesthetic tradition of music's autonomy:

"Scruton's approach relies on taking the first step of claiming that musical events are not tied to the physical circumstances of the real world but separated from them, and capable of 'behaving' in ways that are not constrained by the real world. His notion of movement, gesture, etc. is necessarily abstract and (in some sense) idealised - because the 'things' that are moving are metaphysical rather than physical. The movement and space are metaphorical because the properties of real space and movement have been transferred across to another domain where they have no literal application. (...) [H]e draws a clear line between the tangible and practical world of sound, and the abstract and incorporeal domain of tone - and in doing so aligns himself unambiguously with the aesthetic tradition of music's autonomy."

(Clarke, 2000:217)

Hence, it seems that the metaphysical legacy of the concepts of 'meaning' and 'expression' obscures, once again, one's sight of the required explanation of the musical phenomenon.

Nevertheless, Scruton's approach remains accurate and important because it contributes to restoring the musical surface as the place where the musical meaning happens (i.e., note in the underlying systems, or the deep structure) and is expressed through the movement we hear in and through the music. His insight into Meyer's work, for instance, makes clear what in Meyer's theory remains mysterious. Meyer's account of the listener's expectations remains mysterious in what concerns the real motivations of the listeners to attend the different structural musical elements which make part of the musical surface.[16]

Expectations are just a part of the musical process through which meaning is conveyed. As Scruton (1997:332) observes, "the sense of closure in music is not the primitive fact, as Meyer would have it; on the contrary, it derives from the experience of movement, which cannot be used to explain." Evidence has shown (Fowler, 1986) that sound can specify movement, and sound's grouping is a listener's capacity of detecting real or virtual sources of movement (Bregman, 1990:469; and Clarke, 2001:221). Thus, Scruton's claiming of a musical meaning that is not dependent, like in Meyer's theory, upon an understanding of a set of rules that would provide music's organisation and explain musical content, but dependent on a 'spontaneous result of an imaginative act of attention', seems accurate:

> "Whatever rules are proposed - whether linear, hierarchical, or epistemic - they will misrepresent the organisation of the musical Gestalt, which is a spontaneous result of an imaginative act of attention. The formal relations that we perceive in music neither are, nor result from, a structure below the surface. Form and structure in music are purely phenomenal, and even if our grasp of them can be improved, by an 'emendation of the intentional understanding', the result is an experience of the same kind as the one which stood amended: an experience of movement, life and gesture, reaching through the imagined space of music. Undoubtedly Meyer is right, in arguing that melodic and rhythmic organisation are many-layered, and that the phrase-structure of music is far more complex than it might at first appear. But this phrase-structure is given to the musical ear, and our difficulties arise not in hearing it, but in attempting to describe what we hear in terms which capture its complexity."
>
> (Scruton, 1997:333)

---

[16] This is probably the reason why Meyer's later theory of meaning became more and more 'emotionless', i.e., lesser and lesser emotion-worried or emotion-care. As Cook (2001:59) remarks: "There is something symbolic about the fact that Meyer's first book is the only one he wrote specifically about musical emotion and meaning, the others all focusing on issues of structure and style."

The consideration of music's motion character is surely one of the most urgent claims to be explored concerning the expressive power of music. A music perceptual theory that does not concern music's motion character as an indispensable constituent of musical meaning is perhaps condemned to be unsuccessful. Scruton is surely right when he claims that the organisation of music is to be perceived not as merely movement, but as a gesture or action. He claims, for instance, that:

> "Musical activity is not just movement, but the peculiar form of movement that we call action - the confluence of life and rational agency which distinguishes humanity from every other phenomenon in the natural world."
>
> (Scruton, 1997:333)

And this explains, according to him, the peculiar effect of silence in music:

> "We hear silence as 'Schweigen', a being-silent. It is not cessation of action, but action of another kind - refraining, withholding, refusing."
>
> (Scruton, 1997:333)

Thus, Scruton (1997:365) is right when he claims that "we hear musical movement as action, and not just as movement." But, what is flawed about Scruton's approach is the claim that the perception of movement in music involves the separation between the 'tangible and practical world of sound', and the 'abstract and incorporeal domain of tone'. According to his own words:

> "The activity which animates the musical surface is that which animates you and me - although transferred to another and inaccessible realm, the realm of pure sound, where only incorporeal creatures live and breathe."
>
> (Scruton, 1997:333)

In other words, sound not only specifies the 'rational agency' of the movement but also the bodily action that produces that movement, even if the intention and the nature of what is acting is not always clear enough to be defined through words. According to Clarke (2001):

> "The sounds of music can and obviously do specify objects and events in the world (instruments and the people who play them), and kinds of action, even when the nature of what is acting is unclear or uncertain (we may hear blowing or scraping without knowing exactly what is being blown or scraped)."
>
> (Clarke, 2001:217)

The relationship between the sounds of music and the gestures and actions of the people who plays them, will be further discussed in item 1.7 within the exposition of the 'Ecological Contribution'. A remaining question, however, will be left open, in order to be later discussed: "Why have so many theorists failed to acknowledge that musical movement is, among other things, human movement? (Shove & Repp, 1995:58)." Again the 'heavy' movement of metaphysics can be heard in the wings beats of the idealistic legacy.

1.4   THE STRUCTURALIST CONTRIBUTION

Another commonly shared idea in the field of music and music performance studies considers and assumes 'musical expression' as being directly connected with the musical structure, as being generated by it (Clynes, 1986; Sundberg, 1988; Clarke, 1988; Sloboda, 1988; Rink, 1990; Todd, 1985). The concept that 'structure' functions as a kind of mediator between the concepts of 'meaning' and 'expression', and its appearance as an operating tool in musical research has been made mainly under the major influence of Chomsky's (1957) generative linguistic theory.[17]

According to Sloboda (1985:17):

> "As a speaker is able to produce grammatical sentences precisely because he is able to represent each sentence as an unified structure in which the parts have the kinds of relationships between them which are exemplified in deep-structure tree. Similarly, a composer is also able to produce a masterwork because, as Schenker would claim, he has the intuition of the Ursatz underlying it, which guides and unifies the process of generating the individual notes."
>
> (Sloboda, 1985:17)

---

[17] For a further explanation of the generative approach to music/performance see Schenker (1935), Lashley (1951), Lerdahl & Jackendoff's (1983), Sloboda (1988) Clarke (1995). According to Sloboda (1985), the book by Lerdahl & Jackendoff (1983) suggesting that tonal music takes on meaning and is understood by the listener by virtue of a known tonal grammar, remains the paradigmatic expression of the received view on representation of tonal music. "Lerdhal & Jackendoff's main achievement was to deliver a set of specific operations which, if applied to a piece of tonal music, would yield a complete description of that music, in which each event was marked as being more or less prominent in one of several hierarchical systems. Such descriptions yielded strong empirical predictions about perception and memory (e.g. listeners would recall prominent events more readily than subsidiary ones, would hear prominent events as more important or accented, would fear events as subsumed by more stable ones, etc.)" (Sloboda, 1998:23).

The Chomsky-Schenker parallel, the so-called generative approach, can be further observed at work in musical performance analysis, where performance might be regarded as "a serially ordered temporal phenomenon controlled by a hierarchically organised representation" (Lashley, 1951). By re-examining its core principles, Clarke (1995) comments that:

> "At its heart is the idea that expression comprises systematic patterns of deviation from the 'neutral' information given in the score, which take the form of rule-based transformations of canonical values originating in the performer's internal representation of the musical structure."
>
> (Clarke, 1995:22)

As Clarke (1995) notices, this definition, is not without its problems:

> "How does one distinguish between deliberate departures (or transformations) and mere accidents? And what about expressive markings already in the score (accelerandos and ritardandos, crescendos and decrescendos, etc.)? Must we regard corresponding tempo and dynamic changes in the performance as inexpressive simply because they follow such markings?"
>
> (Clarke, 1995:22)

To avoid these problems while retaining the basic principle that expression is a departure from some norm, Desain and Honing (1992) suggest, that:

> "Expression within a unit is defined as the deviations of its parts with respect to the norm set by the unit itself."
>
> (Desain & Honing, 1992:175)

Briefly said, expression in performance is, thus, believed to come from deviations from the score. Nevertheless, whether the weight of such an idea tends more to the general mechanisms of a communicative large-structural orientation or to the more individual, idiosyncratic and unpredictable components of a specific and detailed artistic expression, a very different sense of music signification and of musical meaning becomes manifest. Two examples of recent studies in music performance may clarify what has just been stated:

- Firstly, Cook (1995:105-117) studied two recordings of Furtwängler conducting Beethoven's Ninth Symphony, and presented evidence of a clear relationship between variations in tempo and Schenkerian principles of structural analysis. He also showed that Furtwängler was highly consistent in

modifications of the tempo despite widespread criticism from contemporary music critics that his timings were erratic and arbitrary.

- Secondly, Shaffer (1992:263-278) analysed the data of four pianists, each of whom performed the piece of Beethoven's Wo0 60 several times, in terms of the unity and diversity of interpretations. By considering the most contrasting performances, he came to the conclusion that what distinguishes them is not a difference in their structural implications, but rather a difference of characterisation.

Shaffer (1992) regards expression, thus, as a means of conveying mood and narrative in music, and can only partially be reduced to a conventional notion of structure. According to these ideas we can have totally different levels of awareness of the music/performance event: it will be perceived as a structural event, even if the perception is only of performance features; or, as an expressive event, even if reported only as structural features. If, then, on one side, it is the large-scale gestures that show the performer's comprehension of a work as a whole - for instance, according to Sloboda, (1985:83) long-term variations are often dictated by large-scale structure of the work; on the other side, smaller but essentially contrasting expressive gestures within a performance may reveal the originality and the idiosyncrasy of a specific interpretation. In Shaffer's words (1992):

"Expression, including the choice of tempo, seems relevant to conveying mood as an aspect of musical meaning. If we further suppose that music can convey an abstract narrative, then we can think of the musical structure as describing an implicit event, and the gestures of musical expression as corresponding to the emotional gestures of an implicit protagonist who witnesses or participates in the event. Thus, the performer's interpretation can be viewed as helping to define the character of the protagonist... Using this conception as a way of relating structure and mood in the music also allows us to see that, in principle, a performer can be faithful to its structure and at the same time have the freedom to shape its moods... If the structures in the music particularly those governing tension and relaxation, define the implicit event, then structure should be the primary determinant of the patterning of expressive gesture over the musical surface. On the other hand, the shaping of expression and the choice of expressive gestures – timing, dynamics, timbre and articulation – is a function of the musical character, and is, at least partly, created by the performer. The concept of an underlying narrative takes the

study of musical expression beyond recent studies that relate expression only to local properties
of the music structure..."

(Shaffer, 1992:265)

Thus, what seems undeniable is the clear tripartite correlation between *musical expression,*
musical structure, and *musical meaning.* However, one should be aware that establishing a
correlation between them is not the same thing as establishing causality. Where do
expressive variations originate and what generates them? According to Persson et al. (1992,
cited in Gabrielsson 1999:503), there is evidence suggesting that, when generating
expression in a performance of unknown pieces, the representations the performer realises
of them may take the form of mental images, characters, moods and events. The expressive
variations may be, then, the physical expression of an expressive intention used to convey
emotion, mood, narrative or any other event, including structural ones. So, it might be that
the perspective approaching expression in music and in music performance as generated by
systematic patterns of deviation from the 'neutral' structure, could be reducing the
understanding of musical meaning to a kind of 'dry mechanicism'. Of course one has to
agree with Cook (1987:229) that:

"The basic thing about an analysis then, isn't what it manages to fit in... but what it manages to
leave out. Schenkerian analysis for instance has been criticised for omitting rhythm: but isn't
the point of Schenkerian analysis that it clarifies the music's structure precisely through this and
other omissions? So an analysis should not aim to be a carbon-copy of the listener's experience:
rather it should simplify, clarify and illuminate it."

(Cook, 1987:229)

Structural analysis is certainly able to accomplish these aims. It assumes that meaning and
signification are contained within a particular piece's structural relations. The musical
object is viewed as an empirical entity to be probed by the subject in such a way as to
reveal its underlying systems or coherence upon which its perception as being
'meaningful' rests. According to structuralism the most stringent conception would be that
"structure is self-regulating and it makes no appeal beyond itself in order to validate its
transformational procedures. The transformations act to maintain and underwrite the
intrinsic laws which bring them about, and to 'seal off' the system from reference to other
systems" (Hawkes, 1977:16). Thus, in a sense, structuralism joins the conception of
musical meaning as immanent and intrinsic to the musical work and, of course, the tacit

dualism of this tradition, in what refers to the nature of the musical object and its meaning. Ruwet (1967), for instance, claimed that:

> "Musical meaning appears only in the description of music itself (...) The signified (the intelligible or translatable aspect of the sign) is, in music, given in the description of the signifier (the material aspect). The only way to have access to the study of the meaning is through the study of the form of the musical syntax, and through the description of the material aspect of music at every level where there is reality."

> (Ruwet, 1967:91)

Structuralism's approach to musical meaning as a sign-system is, thus, based in the 'objectification' of the musical experience and on the consequent subject-object distinction, and, of course, in the denial of the relevance of the subject's psychological and processual involvement in the musical performance, perception and analysis. Nevertheless, the importance of this approach can be seen in the relevance given to the fact that meaning in music "cannot reside in the isolated elements which enter into its composition, but only in the way those elements are combined" (Lévi-Strauss, 1972:210). Structural and significant through the internal organisation of the musical work, meaning in music remains immanent and intrinsic to the music itself. But, if it is simply and exclusively intra-musical, how is it possible that music might determine or influence individual consciousness (the perception or induction of emotion, mood, and movement) and how it might therefore exercise power over individuals (individual and collective ecstasy and hypnosis, for instance)?

Despite all the effort invested by researchers to explain musical meaning through the internal organisation of the musical work relationships, it seems that the formalist and structuralist approaches allow at a certain point a less intrinsic conception of musical meaning. It is the case of Meyer with, for instance, the already exposed 'ethetic relationships', and it seems also to be the case of Ruwet's (1972) account of musical meaning, in a moment where he clearly invokes the character necessarily extrinsic of it within the relation of homology between musical structures and structures of the external reality:

> "The linguistics, structuralists, and generativists, taught us that the internal study of a system comes first with relation to the study of its psychological and physiological conditions. (...) The analysis (of a musical fragment, a piece, etc...) should be able to detect musical structures that

are homologue with other important structures of the external reality or of the lived experience; it is within this relation of homology that the meaning of a musical work can be revealed. For instance: a tonal fragment of music composed of two different parts A and A'; A ends with a suspended cadence, A' starts the same way A does but ends with a perfect cadence. Within the tonal system, we can clearly say that the first part might be interpreted as a movement taken to a certain point and then interrupted or suspended, and the second as a 'reprise' of the same movement, this time taken to the very end. We can see that, in this case, the simple description allows us to reveal a certain structure – (...) – homologue with an infinite ensemble of other structures that one can find in the extra-musical reality or in everyday lived experience.

(Ruwet, 1972:13-14)

Accepting the fact that musical meaning might be revealed through the homology between a formal structure and a lived experience of a structure, Ruwet introduces an extrinsic dimension in his analysis of the musical fragment. He seems to be comparing the musical form to an 'affect' or any other state of 'body awareness'. Thus, Ruwet's conception of 'homology' points to a sort of confusion between what Nattiez (1975:140) calls the 'neutral level' of analysis (the pure formal analysis) with the 'esthesic level' of perception (the public's perception of the musical work):[18]

"When Ruwet says that the interrupted cadence is 'a suspended movement', he makes a semantic judgement that borrows its terms from the extra-musical world... A problem that all authors that try the intrinsic description seem ready to forget."

(Nattiez, 1975:140)

It seems, though, as referred above, that there is more in this idea of 'homology' than might be expected at first sight. Lévi-Strauss (1971:5), for instance, adds to the subject that "the musical work furnishes a grid of signification, a matrix of relationships which filters and organises lived experience; it substitutes for experience and produces the pleasurable illusion that contradictions can be overcome, and difficulties resolved." Again, it makes one think of Ruwet's referred example of homology: "I-V-VI, I-V-I". Corroborating this idea, Shepherd and Wicke (1997:44) claim that some principles of structuralism, enunciated by

---

[18] See later on in this chapter 1.5 "The Symbolic Contribution" the distinction between the neutral level, the poietic level, and the esthesic level within the theory of Nattiez (1975).

Saussure (1966), can be usefully applied to understanding music, and seem to have a special relevance for understanding its processes of signification:

> "Saussure's work remains instructive over and above establishing principles of structuralism of clear relevance to understanding some important characteristics of music as a signifying system. (...) Our position (is) that music cannot, on the one hand, be reduced to the condition of simply reflecting social processes or, on the other, to being a consequence of individual creativity alone. (...) Saussure's work opens up the possibility that music might in some way 'produce' meaning. This possibility makes feasible a consideration of how music might determine or influence individual consciousness and how it might therefore exercise power over individuals."
>
> (Shepherd & Wicke, 1997:48)

Nevertheless, meaning remains with structuralism intrinsically connected to musical structure and immanent to it. Unable to relate accurately the musical meaning with the extrinsic world of objects and events, structuralism leaves us, nevertheless, with the value of a new concept, the 'structure', which is able to defend the integrity of music as a distinct and particular form of human expression and communication.

1.5   THE SYMBOLIC CONTRIBUTION

The symbolic approach searches, contrarily to structuralism, to establish a clear connection between intra-musical and extra-musical meanings by introducing the concept of 'reference'. Arbitrarily conceived within Saussure's (1966) theory, or motivated within Peirce's (1931-8) conception of semiotics, the relationship between symbol (or/and sign) and the referent (object of the 'world' or/and of context) is an important issue to both theories. The term 'reference' joins, thus, the terms - 'expression', 'meaning', and 'structure' - and becomes by this fact, another element contributing to the reflection and examination of expressiveness in music/ performance. The symbolic contribution to this issue is, thus, devoted to explore how music may function as an instance of 'semiosis'.[19]

---

[19] According to Peirce (31-8: 1906, Pragmatism, v), "semiosis is any dynamic action or influence physical or psychological that involves always the co-operation of three subjects, such as a sign, its object or referent, and its interpretant. According to Greimas & Courtés (Semiotique, 1979:339), "semiosis is any action that by establishing a relation of reciprocal assumption between the expression of the form and that of the content, in Hjelmslev (1971) terminology – or between signifier and signified (Saussure, 1966) – creates signs: in this sense, any action of language, for instance, establishes a semiosis."

According to Bowman (1998):

> "The terms 'reference' and 'expression' are both subject to abundant confusion in discourse
> about art and music. Reference is often aligned with language and expression with feeling.
> Reference is often conceived as a transitive or mediating function, to which expression's
> immediacy is presumed opposite. As well, expression is often construed as a kind of emotional
> arousal, while reference is cool and calculated."

(Bowman, 1998:227)

One of the best known explanations of music symbolism is, as stated above, Langer's presentational symbol theory, which is based on the iconic resemblance that stresses the isomorphic relationship between the musical symbol and the form of the feeling expressed by it. Langer's music symbol is, as said, presentational, i.e., without assigned reference.[20] Because Langer's theory has been enough commented (see pg. 13, item 1.2 "The Iconic contribution"), the exposition will turn now to other authors.

Francès' (1958) 'musical symbolism' considers that music is able to symbolise the space, the movement inside the space and, even, its affects. According to him, "if music is able to induce emotion it is because, first of all, music is able to induce movement" (Francès, 1958:342). It is well known that there is a strong relation between motion and emotion. "Almost every emotion stirs up a movement" (Goblot, 1901:69). Music "moves people to tears" (Sloboda, 1998). The word 'emotion' itself comes from the Latin word which means 'move' or 'stir up' (Reber, 1985). "Musical movement is human movement" (Shove and Repp, 1995). Reflecting on this interface of music and movement, Francès (1958) presents the following conception:

> "The similarity between rhythmic and melodic schemes and the gesture schemes that are
> present in human behaviour, represent one of the prominent elements of music expressive
> language... The fundamental psychic states (quietness, excitement, tension, relaxation,

---

[20] Langer's theory can be expressed, as said, by Pratt's slogan – "music sounds the way emotion feels" - i.e., the music shows us the appearance of feeling in a perceptible symbolic projection, but it does not refer to a public object, such as a concrete and discrete emotion or feeling. The choice of presenting Langer's theory in a different and previous item 1.2 "The Iconic Contribution" is connected with the fluency and clarity of the exposition, but essentially with the fact that, Langer conceives a theory where the symbol is, in contradiction with the general symbolic theories, without assigned reference. According to (Bowman, 1998:204) "it is an well-accepted idea among contemporary semiotic theories that both symbols and signs are instances of referring".

depression) are usually translated by forms of gestures with a given rhythm, tendencies an spatial directionality (obstinate repetition, diversity, periodicity, evolution)... The transposition of these rhythms, tendencies, modalities of movement to the sonorous plan constitutes the foundation of music's expressive language."

(Francès, 1958:299)

So, first, it seems that through the internal properties of sound and the intra-musical relationships between them, music is capable of describing movement by revealing its acoustical properties. And, second, it also seems that through this movement music is able to reveal, as Shepherd and Wicke (1997:127) affirm, "the inner, physiological life of individuals in terms of the way the internal configurations, textures and movements of their bodies affect (and are affected by) the quality of sound production." Imberty (1975) resumes Francès thought as follows:

"During a musical audition a strong and significant relationship can be established between the bodily manifestations and the emotional state of the individual. And that, in such a way that when one hears music that evokes either a bodily scheme or a psychological attitude or state, the experience of the referred relationship will appear in the verbal explanation of the music in the form of responses that report movement or affect."

(Imberty, 1975:95)

Thus, it seems that the musical symbolism of Francès recognises the power of the music to refer to the inner processes of subjectivity and, also, to the more extroverted bodily movements. Margolis (1987:228), goes even further, and conceives musical meanings as "deeply embedded in the various and ever-evolving social practices that give rise to them." Nevertheless, and according to him, the multiplicity and contingency of musical meanings do not allow the semiotics of music to become a systematically oriented science like it is the case with the semiotics of language. Sharing a different opinion, Nattiez (1990) believes that a post-structuralist semiology, one which aims not to explain music according to scientific rules of communication, but one which is capable of developing and of producing good results by describing a larger range of potential meanings inherent to it, is possible:

"The musical work is not merely what we used to call the 'text': it is not merely a whole composed of structures...Rather, the work is also constituted by procedures that have engendered it...and the procedures to which it gives rise: acts of interpretation and perception."

(Nattiez, 1990:ix)

Nattiez is aware of the fact that according to Saussure's structuralism the relationship signifier-signified has to be stable and invariant, that is an arbitrary and static relationship. So, instead of taking into consideration the arbitrary and static relationship of Saussure's sign relationship, Nattiez draws his account of 'referring' from Peirce's theory (1931-8). According to Peirce's (1931-8) semiotics, a sign, or *representamen*, is something that becomes connected to something else (another sign or its 'object'), the referent, in such a way that it draws a third element, the interpretant, into a relationship with the same 'referent'. But, as Nattiez (1990:7) explains "the process of referring effected by the sign is *infinite*." For the interpretant also may function as a sign, mediated by other potential interpretants that relate to the referent in still different ways, and so forth, ad infinitum. So, instead of the linear and unidimensional relationship 'signifier-signified', what Nattiez proposes is a concept of 'sign' which engenders rich and multiple meaning connections. The importance of this statement for the world of music and music performance will be fully understood if one thinks that the 'object' to which performance refers does not exist except within and through the infinite web of 'interpretants'. As Nattiez (1990) explains:

> "The object referred to by a sign is always virtual, it does not exist except within and through the infinite multiplicity of interpretants, by means of which the person *using* the sign seeks to *allude* to the object. (...) What signs refer to is always contained within the *lived experience* of those who use them."
>
> (Nattiez, 1990:7-8)

So, it is this 'lived experience' and the multitude of interpretants that interface with it that brings a totally different understanding of musical meaning. According to Nattiez (1990), the study of musical meaning should not be just concerned with revealing invariant structural bases or rules of communication. Meaning in music is engendered at the intersection of the structural relationships of the sonorous object and the lived experience of those who placed themselves in relation to it.

Thus, contrary to the easy and successful process of language communication, the process of communication within music is far more complex and consists, according to Nattiez's semiological approach, in the interaction of three different levels. Bowman (1998) resumes it as follows:

"The interaction between the producer and the object is a constructive process, one in which, something material is brought into existence by someone who produces it. This Nattiez (after Molino) calls the poietic process. What is thus created has a level of materiality, an existence that bears the "traces" of the poietic process that brought it into being: the immanent (or alternatively, the neutral) level. Where the communication model errs is presuming that the trace is transparent, or that it serves as a conduit for transmitting the meaning or intentions of its producer. On the contrary, the relation of the perceiver to the trace is, like the producer's, a creative one. This act of apprehending, interpreting, and giving meaning to the trace, of bringing the trace into various possible relations with the perceiver's horizon, Nattiez calls the esthesic process. (...) Poietic orientations locate 'music' in generative or reproductive acts, regarding it as an art concerned with the creation and manipulation of sonorous material. Immanent accounts gravitate toward music's 'intrinsic' features, the sound object's intramusical relations and sonorous properties. Esthesic orientations see perceptual pleasure and displeasure as music's definitive characteristic, focusing on the way music is experienced."

(Bowman, 1998:242-246)

Nattiez's approach takes as a source not only Peirce's triangle – sign/interpretant/referent – but also the tripartite semiological scheme of Molino (1978) which involves 'producers', a 'symbolic form' and 'receivers'. Production or creation results, in Nattiez's version, in the 'poietic dimension'; reception accounts for the 'esthesic' dimension; and the symbolic form, or neutral level, a concept inhered from the structuralism, manifests a level of internal organisation, that has to be analysed. But, contrarily to structuralism, Nattiez does not presume that the immanent level is sufficient in accounting for music's meaning:

"Symbols manifest a level of specific organisation, that has to be described. But this level is not sufficient: the poietic lurks under the surface of the immanent; the immanent is the springboard for the esthesic. (...) Musical analysis consists in a recurrent dialectical oscillation among the three dimensions of the object."

(Nattiez, 1990:29-32)

So, according to him, musical meaning is not to be investigated at the level of immanent structures but in the "behaviours associated with sound phenomena" (Nattiez, 1990:65). Nattiez's approach is a decisive renunciation of the structuralist and formalist point of view that considers musical meaning as contained in the intrinsic relationships of the sonorous

configurations within the sonorous object. Nevertheless, he conceives musical meaning as being engendered within a 'symbolic system' with the potential to refer to a multitude of interpretative experiences. Meaning in music is directly connected to the basic act of referring that the 'musical work's material presence' (the neutral level) gives rise to and, by the same gesture, is implicated in the infinite webs of 'poietic' and 'esthesic interpretants', giving access to multiple internal and external sorts of referring. This concept of musical meaning will give, thus, access to the social and personal lived worlds directly implied within the analytical process for any given work, as well as to the intrinsic and absolute internal worlds of all the musical-like actions. As Bowman (1998) explains:

> "The meanings the 'musical symbol' engenders are not structurally determined, but are instead, social constructions whose 'musicality' is always relative to cultural habit and the way people uses them. If this sounds more like anthropology or sociology than semiology, we have to recall that in Nattiez's view what is distinctive about semiology is its persistent effort to elucidate the relationships between the constructive and functional features of music and its materiality. (...)
> An adequate account of musical meaning in his view must accommodate all possible forms of musical referring. The intrinsic or internal referring of the aesthetic/absolutist tradition is one such form. It embraces both intramusical referring – the syntactical/implicative relationships which Leonard Meyer and others have explored in such detail – and intermusical referring, the situation of particular musics within certain styles and conventions. Extrinsic referring, that which is semantic in character implicates experience that is spatiotemporal, kinetic, and broadly affective in character. But 'introversive' and 'extroversive' semioses, as he calls them, are not discrete processes or domains. They are both processes of referring, and in actual musical experience they are always tightly intertwined. Indeed one of the most distinctive qualitative features of much musical experience is the state of instability and dynamic fluidity that exists between these two modes of referring."

<div align="right">(Bowman, 1998:248)</div>

However, as pluralistic, anti-dogmatic and open-ended this 'semiological' project might seem to assume the investigation of musical meaning to be, it is not uncontroversial or unproblematic in the way it tries to accommodate and explain the variety and multitude of 'interpretants' engendered by the musical symbol. In fact, one of the criticisms of Nattiez's work comes from Shepherd and Wicke (1997), who believe that Nattiez's notion of 'immanence' retains strong resonances of traditional positions on questions of meaning in

music. According to Nattiez (1990:12), music, embodied physically and materially in its symbolic form (the neutral level) seems to carry a 'trace' of intended 'poietic' meaning:[21]

"The symbolic form is embodied physically and materially in the form of a trace accessible to the five senses. We employ the word trace because the poetic process cannot immediately be read within its lineaments, since the esthesic process (if it is in part determined by the trace) is heavily dependent upon the lived experience of the 'receiver'. Molino proposed the name neutral level... or material level... for this trace."

(Nattiez, 1990:12)

So, the 'intended meaning' (for example, the thematic material in which the musical work is composed) allows for the establishment of an explicit symbolic relationship between the immanent musical structure and the poietic gesture. According to Nattiez (1975:59), "the musical semiology is possible only when 'poietic' and 'esthesic' phenomena find a 'trace' within the musical object." And it should stop "when it is no longer possible to establish convincingly explicit relationships between the immanent 'structures' of the musical object and the esthesic and poietic interpretants." Thus, on one side, the concept of immanence connected to the neutral level seems to depend not only in the used musical materials, but also on a sort of material 'print' of the poietic intentionality, i.e., the use and the intention the creator of the musical work gave to these materials. On the other side (the receiver's side), meaning seems to be constructed not only according to the qualities inherent to the neutral level, but also according to the 'remaining' immanent poietic intentions. Shepherd & Wicke (1997:19) explain that:

"In one sense Nattiez's work possesses the potential to dispense with the notion of immanent meaning in music (at the 'level' of the 'neutral') by allowing for the construction of meaning on the part of the 'receiver' in a manner related to but distinct from this 'niveau neutre' (because this process of meaning construction is 'heavily dependent upon the lived experience of the receiver', and only 'in part determined by the trace'). Yet in another sense Nattiez retains a certain notion of immanence (the trace) in quite rightly refusing to allow that just any meaning can be passed through and assigned to musical materials themselves. However, this notion of

---

[21] In fact, according to Nattiez (1975:50), one of the principal characteristics of the *neutral level* "is to serve as an anchorage to the 'poietic' and 'esthesic' approaches." However, according to Nattiez (1975:59), what is meant by 'intended poietic meaning' is only the author's "*realised intentions* (for instance, the thematic material in which the musical work is composed)".

immanence seems to depend, not on the musical materials themselves, but on the perpetuation within them of a trace of the poietic dimension responsible for their configuration, a trace which then proceeds in part to determine the esthesic dimension. Although meaning is constituted by the 'receiver', it is constructed – it would seem – not simply according to the possibilities inherent in the 'niveau neutre', but in addition according to the trace of the poietic dimension which remains immanent at this level. A careful reading of Nattiez reveals that the notion of the 'niveau neutre' is handled ambiguously - (...) This problem with the 'niveau neutre' occurs precisely because Nattiez is unclear on the question of the relations of music as a process of creation and reception. Nattiez's work as a consequence throws little light on the question of resolving the tension between the opposed poles of the arbitrary and the immanent."

(Shepherd & Wicke, 1997: 20)

What stands as fundamental to this criticism is the idea that music should be seen as a social and cultural constituted signifying practice and as a particular and irreducible form of human knowledge, expression and communication. In Shepherd & Wicke's (1977:2) own words, music should be considered:

"As a distinctive and irreducible signifying practice (with) a structure with features distinct from those of language, as well as (...) communicating in ways which are quite different from those of language."

(Shepherd & Wicke, 1997:2)

So, sounds in music are not substitutes for sounds of language or, even, able to be substituted by sounds of language. Sounds of music are to be understood as being capable of conveying and evoking powerfully, yet imprecisely enough (comparing to language), a set of different experiences and effects. Sounds of music are to be understood idiomatically and idiosyncratically. That is the reason why the multidimensional and polyvalent character of the experiences evoked by music is practically impossible to be organised and articulated in the linear manner of language. According to Feld (1984:14), this is the reason why people find it difficult to articulate thoughts when discussing the feelings and the effects they experience by listening to music:

"When people say, 'It's different from...', 'It's a kind of...', 'It sort of reminds me of...', and things of this sort, they are creating locational, categorical, and associational features. They are caught in a moment of interpretative time, trying to force awareness to words. They are telling us how much they assume that we understand exactly what they are experiencing. We take it as

socially typical that people talk this way about music, stringing together expressives, and we assume that this confirms what we are all supposed to know that at some level, one just cannot say with words what music says without."

(Feld, 1984:14:)

Nevertheless, the multidimensional and polyvalent nature of musical meaning does not imply that music becomes an 'empty sign' (Kristeva, 1989:309) where all meanings are possible, and this because, no concrete and discrete meaning is possible through musical sounds. Bierwisch (1979), for instance, has the opinion that though musical sounds may be related to different and variable contents, it does not mean that just any meaning can be attributed to them:

"This relationship (musical signs and non-musical factors) can be paraphrased in terms of two apparently opposing statements: it is practically impossible to fix the thoughts or notions which are evoked by music; and, the connections between musical configurations and non-musical thoughts and notions are by no means arbitrary."

(Bierwisch, 1979:57)

The major criticism that Shepherd and Wicke (1997) make of the post-structuralist approach to musical meaning has, in fact, to do with this status of 'no man's land of meaning' that post-structuralism attributes to music. If, music's sounds (or as they put it, the 'sonic event') were capable of all meanings by reason of its status as an 'empty sign', they could never be implicated in the meanings they engender. Kristeva (1989), a post-structuralist author, claims, for instance, that music is, in the last instance, reducible to the condition of the pre-symbolic, the pre-linguistic and the pre-discursive communication, because music's materiality does not seem capable of guaranteeing a relative autonomy as a signifying practice:

"It (music) does transmit a 'message' between a subject and an addressee, but it is hard to say that it communicates a precise meaning. It is a combinatory of differential elements, and evokes an algebraic system more than a discourse. If the addressee hears this combinatory as a sentimental, emotive, patriotic, etc., message, that is the result of a subjective interpretation given within the framework of a cultural system rather than the result of a 'meaning' implicit in the 'message'. (...) the musical code is organised by the arbitrary and cultural difference between various local values: notes."

(Kristeva, 1989:309)

This conception of music is also shared by Barthes (1989), another post-structuralist author, when he theorises music as a world of sound 'outside' that of the language. Music is, according to him, the ideal 'Other', a place where the symbolic order of the Western discourse is absent, and where one has the access to a domain in which one can be exempt from the constraining 'duty' of meaning. He considers that musical meaning exists outside the limits of the world of language, in a domain where, as Middleton (1990:266) put it 'anything goes':

> "Music has a privileged place because it can replace everything else: it is the degree zero of this system: according to the needs of enthusiasm, it replaces and signifies journeys, Women, the other arts, and in a general manner any sensation. Its signifying status, precious above all others, is to produce effects without having to inquire as to their causes, since their causes are inaccessible. Music constitutes a kind of *primal state* of pleasure: it produces a pleasure one always tries to recapture but never to explain; hence, it is the site of a pure effect…an effect severed from and somehow purified of any explicative reason, i.e., ultimately, of any *responsible reason.*"

> (Barthes, 1989:299)

As, Shepherd & Wicke (1997:94) comment: "(according to post-structuralism) either we have language, discourse and reality, or we have the spontaneously unstructured world of ruptured, ecstatic *jouissance*, replete with luscious, Dionysian pleasures". So, it seems that, on one hand, Barthes and Kristeva clearly place the need to accept that sounds of music function as if a direct material link was possible between them and the "somatic pathways of the body" (Shepherd & Wicke, 1997:97). On the other hand, their work seems willing to refuse to understand how this functioning of the sounds is connected to the musical structure and its interaction with people. Thus, it seems that according to them, meaning in music is 'emptied' through the production of an ideal 'Other' to the world of the symbolic order of Western discourse, i.e., the semiotic, the women, the foreign, the intangible. As Kristeva (1989) argues:

> "Music takes us to the limit of the system of signs. Music is a system of differences that is not a system that means something, as it is the case with most of the structures of verbal language. Music therefore has a trans-linguistic status (…) it is an 'empty' sign."

> (Kristeva, 1989:309)

Nevertheless, Kristeva's work provides an explanation that makes possible a linkage between sounds, sounds of music, and the human body. As Shepherd & Wicke (1997:80) notice, "sounds, together with rhythms, gestures and colours, form for Kristeva one of the means through which the energy drives linking the bodies of the mother and child in one sensate universe are channelled and directed". These statements have a strong resonance with the results of the development research and infant studies of Stern (1985), Papousek (1996) and Trevarthen (1999), according to which behaviour of children is based in a *musical* interplay between mother and child. Infants imitate the vocalisations, facial expressions, and gestures of others around them and this gradually integration of the sensory-motor mechanisms of the body into higher levels of concept formation and metaphorical discourse is what has been referred as the notion of *embodied meaning* (Lakoff and Johnson, 1980). Sounds, thus, and from infanthood on, provide apprenticeship for an awareness of the world and a connection with others. According to Shepherd & Wicke (1997):

> "Kristeva's work provides thus the 'concrete connection' lacking in post-Freudian accounts of the relations between music and the unconscious, as well as a sense of fundamental ordering in sound which lies *unambiguously* outside the realm of the semantic... (pointing) to the possibility of a music *structured according to a-semantic principles* whose sounds are capable of affecting directly the materiality of the human body."
>
> (Shepherd & Wicke, 1997:81)

In this special sense, the work of Barthes (1985:308) is also concerned with the music connecting directly to the body and as a signifying practice lying outside the world of signs, meaning and signification:

> "Music is a field of *signifying* and not a system of signs, the referent... is the body. The body passes into music without any relay but the signifier. (...) the body in a state of music."
>
> (Barthes, 1985:308)

Thus, it seems that, at least according to some of the post-structuralism accounts, it is the body that musical sounds materiality primarily address and link. Rosolato (1974), for instance, considers the fact that music originates in the primacy of the mother's voice for the new-born infant. According to Shepherd & Wicke (1997:88), most post-structuralist

theories conceive the body as constituting the material grounds and pathways for the development, investment and experiencing of emotional energy and affective states. However, the problem with this way of thinking is, according to them, that, on one hand:

> "it assumes that the body, (...), displays no logics or potentials for signification which can stand independently (although relatedly) alongside those of the language. It is assumed that the body can only take on significance which can stand in relation to the development and maintenance of identity by being interpellated in the universe of language."

Shepherd & Wicke (1997:88)

And, on the other hand, (Shepherd & Wicke, 1997:87), that:

> "music, even in its relevance to the preconscious and conscious, is reduced ultimately to the condition of the pre-symbolic, the pre-linguistic, the pre-discursive and the unconscious. (...) Music's materiality does not seem capable of guaranteeing it a relative autonomy as a signifying practice distinct both from society and the individual, in this case to the pre-linguistic and the pre-symbolic subject."

Shepherd & Wicke (1997:87)

It seems, thus, that these theories, despite the importance they attribute to the body and to music's materiality, do not allow music to be conceived of as a signifying practice with its own specific characteristics and modes of signification, giving direct access to knowledge and experience, independent of language. Language is, thus, conceived within that theoretical framework, as the only mode of communication responsible for signification. Music's materiality is confined to comprehend just the material base of the signifier, devoid of the signified. It is, as Kristeva argued, *significance* in its purest form, the *semiotic*, i.e., the pre-linguistic and pre-symbolic subject. The *semiotic*, as Lechte (1994) explains, is sound, rhythm and movement anterior to sense, and the multiple bases of enunciation linked closely to drives and impulses.

Shepherd and Wicke's theory draws upon Kristeva and Barthes' accounts of a world of sound 'outside' that of language, to the limits of semiology of language, and to the threshold of semiology of music. The work of Shepherd and Wicke is based on the conception of music as a new world of sound 'other' than language, and searches to develop a semiology of music based on an understanding of how music functions as a structured and structuring 'medium'.[22] The aim of their semiology is to grasp the processes through which, music articulates significance and affect in a purely structural, but corporeal, manner:

> "We would argue that, as corporeal and somatically manifest, music is both structured and structuring. As such, it resonates powerfully within the lived, corporeal and somatic experience of the listener. To hear a voice, a musical sound, is to 'have knowledge' of the corporeal and somatic state which produced it. The reaction is both sympathetic and empathetic."

> (Shepherd & Wicke, 1997:180)

Empirical evidence from the performance research field seems to corroborate this idea of music being able to articulate significance and affect in a purely structural manner. Sloboda (1998:27) has shown how particular structural musical features impacting on us are capable of affecting directly the materiality of the human body, provoking a reaction whose significance has been 'translated' by the individuals subjected to the experiment into tears, shivers, gooseflesh and racing heart. In other words:

---

[22] Shepherd & Wicke (1997) oppose the concept of 'medium' to Nattiez's concept of 'neutral level'. According to them: "The concept of medium has two distinguishing characteristics. Firstly, it conceptualises the use of sounds in music as being, indeed, of a structural nature, both 'internally' and 'externally'. Secondly, while the 'medium' conceptualises the use of sounds in music as being in this way structured and structuring, it in no way assigns an agency of *achieved meaning* (...). It is as a consequence of this that we believe the concept of the medium to differ significantly from Nattiez's *niveau neutre*. We would argue that music can only be understood to display a *niveau neutre* in the sense of shaping the materiel grounds and potentials for meaning construction, not the processes of meaning construction themselves. (...) Once, meaning in music is taken to be located in processes between the sounds of music and people rather than in the sounds of music themselves, elements of signification can no more easily be disengaged from people's somatic states of awareness than they can from the sounds of music as medium" Shepherd & Wicke (1997:116-120).

"The abstract (visually 'non-denotative' character of the 'internal properties' of the sounds)
(...) means that they *have* to function in a structural manner if they are to function at all in
being *implicated* in processes of meaning construction, if (...) a technology of articulation is to
occur through the circuit of signification..." [23]

(Shepherd & Wicke, 1997:158)

Thus, music, through its structural features, offers the possibility of investing and of being
invested with meaning. This investment is not, however, constituted through a univocal
relationship between music and meaning. It is not, as it is in language, an arbitrary and
conventional relationship between structured sounds and processes of subjectivity. This
one-to-one relationship between musical sounds and meaning does not exist in music. As
Middleton (1990:154) explains:

"It seems likely, too, that in the relationship between (musical) form and... 'experience'... the
two must 'dock' – rather than the one completely producing the other or the conjunction being
purely one of juxtaposition. The docking may be relatively loose; but the parties must meet
within certain *limits of tolerance*."

(Middleton, 1990:154)

As Shepherd & Wicke (1997:178) explain, "not all music can receive all possible
'intended' meanings and not all possible 'intended' meanings can be successfully invested
in all music." Nevertheless, this 'relatively tight iconic intersection' (the sounds of music
and processes of subjectivity) is, according to Shepherd & Wicke, *meaningful*:

---

[23] Shepherd & Wicke (1997:158) described this circle of signification as follows: "Since the sounds of music
start out by acting in a mimetic and so homologous fashion in relation to other sounds (for example the
sounds of the sea), there is nothing to prevent them acting at the same time in an iconic fashion in evoking
more directly the logic and structures of the inner life (the 'first order' states of awareness occurring at the
primary level of signification) to be evoked by second-order states of awareness (occurring at the level of
connotation or secondary signification) such as 'oceanness'. There is, in other words, a circuit rather than a
line of signifying connections involving sounds in music: sounds of the external world, whose invocation
constitutes third-order states of awareness; second-order states of awareness (connotation and secondary
signification); and first-order states of awareness (primary signification). These experiential moments in the
circuit are all mutually reinforcing as the concretization made possible by their structurally homologous 'fits'
moves their connections towards the realm of the conventional and the customary." For further discussion on
this subject see (Shepherd & Wicke, 1997:153-168).

> "As an auditory phenomenon music offers up a material ground in and through which
> individual subjects can invest and map their own meanings, their own experiences. Music is
> meaningful to individuals inasmuch as inherent characteristics of its structures and textures are
> suitable to the investment in them of the meanings – themselves coded and experienced through
> the rhythms, textures and structures of the body - that such characteristics may call forth."

> (Shepherd & Wicke, 1997:178)

Indeed, Baily (1977:330), for instance, argues that "music can be viewed as a product of body movement transduced into sound". According to him, player's movements affect musical structures in an effective way and the motor patterns implicated in different kinds of music change according to the particularities of each music because they are the result of the interaction between the morphology of the instrument and the player's sensorimotor capacities. In fact, as Shepherd and Wicke (1997) explain, the way in which these 'internal states somatically experienced', interrelate and intersect with the 'sounds of music', and become 'source of externally uttered movements', is one of the most exciting sources of contemporary musical performance research:

> "The internal dispositions and configurations of bodies that find expression in externally
> perceptible bodily movement affect fundamentally the dynamics of sound production from the
> most basic and inalienably human of all musical instruments, the human voice."

> (Shepherd & Wicke, 1997:179)

This is especially interesting, since it is the central subject of the research of this thesis: the human voice and how it articulates within singing the investment of emotional meaning. It seems, thus, that an interrelationship can be articulated between: the internal state of the performer, corporeally and somatically experienced, the manifested voice production during the act of singing, and the lived corporeal and somatic experience of the listener that resonates with it. Sundberg (1987), for instance, interprets the relationship between body movements, human emotion, and the physiology of voice production in these terms:

> "There is a close correlation between body movements observable with the naked eye and
> hidden body movements. Examples of normally invisible body movements can be found in
> laryngeal cartilages, most of which are involved in the regulation of voice pitch. If it is true that

a particular pattern of expressive body movements is typical of a specific emotional mode, then we would expect a corresponding pattern of, for example, voice pitch in speech produced in the same emotional mode. In other words, it is likely that expressive body movements are translated into acoustic terms of voice production."

(Sundberg, 1987:154-5)

Middleton (1990) corroborates this idea by explaining that by listening to a voice singing, or an instrument being played, listeners react both sympathetically and empathetically:

"Listeners identify with the motor structure, participating in the gestural patterns, either vicariously, or even physically, through dance or through miming vocal and instrumental performance."

(Middleton, 1990:243)

In fact, studies of motor imagery (Gallese & Goldmann, 1998; Fadiga & Gallese, 1997; Fadiga et al., 1998) report experimental evidence which suggests that understanding the observed behaviour of others involves imagining performing the same or similar actions. Cox's notion (2001:196) of an 'embodied musical meaning'[24] is, also, based on the idea that overt or/and covert mimetic participation is fundamental to musical experience. Baddeley & Logie (1992) suggest that "if comprehension of spoken words ought involves covert imitation it seems reasonable to expect that comprehension of sung words ought to involve covert imitation as well (quoted in Cox, 2001:200)." As Cox (2001:199) writes: "Since human musical performance involves specific motor actions, these studies on mimetic motor imagery become relevant for conceptualisation of music as performed." In fact, enough empirical evidence (Shove & Repp, 1995; Gabrielsson & Juslin, 1996; Todd, 1992, 1995, 1999; Sundberg et al., 1995; Davidson, 1991, 1993, 1994; Clarke and Davidson, 1998; Clarke, 2001) nowadays supports the idea that musical performance involves the production and the recognition of specific motor actions. In this sense, Shepherd & Wicke (1997) observe that, to conceive and understand how the sounds of

---

[24] This notion is of course draw on Lakoff & Johnson's (1980) notion of *embodied meaning*. As Johnson (1987) explains: "The inferential structure of our abstract reasoning is a high refinement upon orderings in our bodily experience, a refinement that ignores much of what goes into our reasoning" (Johnson, 1987:5). See also Jackendoff (1988) and Lidov (1987) for a further discussion on the bodily basis of musical meaning. The short review of empirical evidence in music performance presented above will be reviewed again later, in Chapter 3, in a more detailed manner. Nevertheless, given the present context of the exposition, it seemed helpful to illustrate the theoretical exposition and what is meant by the parallel of structures between the two sets of processes above referred: musical processes and processes of affective states.

music act as a structure when they impact on us and reveal musical structure to us, one should avoid any parallelism with the way the sounds of language act and reveal language structure to us:

> "[I]t is *our* argument that infinite free play of imagination and signifiers in language that is permitted by language's 'semiotic moment' is ultimately constrained by the structural *and* iconic principle according to which music articulates affect and meaning. The principle of *structuring* applies to the external world of 'the actions of language' as put in motion and internalised by people. The principle of *the iconic* applies to the internal world of states of awareness, which gives rise to *and absorbs* elements of signification as an aspect of people's dialectical relations with the external world. It is the abstract yet embodied *structures* of states of awareness, created and maintained as structures through the material iconicity of music's 'semiological moment', which 'play host' to the structures of signifieds and their mutual relations as they absorb them as mental concepts. If the manipulation of the structures of music as forms that are *symbolic* is made possible through language's capacity to denote particulars, then the maintenance of language as a structured and symbolic form *and* its consequent capacity to manipulate particulars is made possible through music's endemic character as an iconic structure."

(Stephen & Wicke, 1997:201)

To conclude, one should add that music does not have to be purely iconic or homologous any more than language has to be purely arbitrary, in the way they seem to signify. As Shepherd & Wicke (1997:209) acknowledged, music has the ability "to go beyond the homologous and the iconic signifying potentials of sound in constituting itself as a performance event through words, images and movement"; the same way, language has the ability 'to go beyond the arbitrary potentials of sound in constituting itself as a performance event through the 'paralinguistics' of rhythm, inflection, image and movement."

1.6   THE PHENOMENOLOGICAL CONTRIBUTION

The preference for a phenomenological methodology, owing to the richness of the perception and the detailed description it affords, has become a critical tool to overcome the difficulties created by the legacy of both empirical and rationalistic approaches. Phenomenology refuses the idealistic and dualistic conception of reality of those traditional

accounts, and believes that between 'apparent' and 'real' there is no metaphysical gap. As Bowman (1998) explains:

> "Phenomenology denies the idealistic splitting of reality into subjective and objective, apparent and actual, and attempts to avoid the reduction of one pole to the other. In actual human experience, phenomenology insists, there is no gulf that requires a metaphysical bridge, no mutually exclusive relationship between knower and known. (...) Appearances, phenomenology maintains, are not illusory or suspicious, but potent sources of foundational knowledge. (...) If theoretical habits, biases, and presuppositions are set aside, the mind is capable of perceiving the essential in appearances."
>
> Bowman (1998:255)

Thus, phenomenology tries to avoid the twisted explanations of reality created by the metaphysical accounts and tries to 'catch' it at a 'vital' and 'experiential' level where distinctions between subjective and objective, apparent and real, perceiver and perceived, are not appropriated. It is a 'experiential world', the world whose reality phenomenology tries to 'catch'. So, when this methodology is applied as an approach to music, its efforts are not concerned with explaining what music is or means. Phenomenological method is not primarily concerned with explaining, but rather in describing. Understanding music's meaning demands a richness and fullness of description - how is music actually heard, perceived, and experienced. As Bowman (1998) puts it:

> "As a philosophical approach to music, the phenomenological typical resists efforts to explain what music is 'about', resembles, symbolises, or is useful for, preferring instead to describe as richly as possible what music itself says, how music is experienced. Instead of explaining, it describes. Thus, phenomenological musical inquiry promises a 'return to beginnings'. Its concern is not so much to establish absolute or universal truths, but to help recover the richness and fullness of the experientially given,"
>
> (Bowman, 1998:255)

This seems to be the reason why the phenomenological method is at the base of some of the more up-to-date approaches to musical performance and of some other 'lived' aesthetic experiences (Stubley, 1966; Burrows, 1980; Clifton, 1983), and also why it seems highly resonant with Johnson's (1987) bodily basis of mind theory, and with Gibson's (1979) ecological perceptual theory.

The phenomenological method is an issue of Husserl's philosophy (1964) and of Merleau-Ponty's (1962) epistemology. According to this new theory of knowledge, perceptual experience gains a foundational epistemological status where the body plays the role as our instrument of engagement in the world. The defining ground of knowledge is not abstract reason, but rather perceptual experience, and the origin of human perceptual experience is, always the body. Merleau-Ponty (1964:13) argues that, "the perceived world is always the presupposed foundation of all rationality (...)". But, according to him, rational or abstract concepts "would not be better known to us if we had no body and no sensibility (...)" (Merlau-Ponty, 1962:394). Thus, our essential way of being in the world is bodily constituted rather than being a conscious experience, and perceptual knowledge is fully constituted in the act of perception, requiring no further contribution from a 'disembodied' intellect. The resonance with Johnson's (1987) theory is strong.[25] According to Sallis (1973):

> "Perception is always spatially situated, is perception from somewhere, so that things inevitably present themselves to me in a one-sided fashion; they offer themselves to my perception by way of profiles, and I never manage to perceive a thing from everywhere simultaneously, i.e., from nowhere."
>
> (Sallis, 1973:36)

But if something is only partially revealed, how can it be nevertheless fully constituted as a result of the act of perceiving? According to Bowman (1998:261), Merleau-Ponty's explanation is that:

> "The appearance of any particular profile is schema-dependent: perception does not consist of apprehending a single profile, but is always attended by awareness of other potential profiles implicit in the operative schema. Consequently, perception always transcends the particularity of a given perspective in the direction of its object. This is what enables recognition of an object's constancy when observed from different profiles, assuring that different profiles, are not mistaken for different objects. (...) What the act of perception adds to mere sensation, then, is a sense of depth beyond surface, a recognition that its object always consists in more than just this one presentation. Perception carries with it a sense of the inexhaustibility of ways its objects can present themselves."
>
> (Bowman, 1998:261)

---

[25] Johnson (1987) states, for instance, that: "The inferential structure of our abstract reasoning is a high refinement upon orderings in our bodily experience, a refinement that ignores much of what goes into our reasoning" (Johnson, 1987:5).

This perception's process is, of course, highly resonant with Gibson's theory of perception and information.[26] Although Merleau-Ponty, as well as Gibson, have not directly applied their theories of perception to music and acoustic perception, the implications and consequences are obvious in both theories. Musical meaning is clearly implicated in the perception of music's experience and performance. As Bowman (1998) stresses:

> "Contrary to the claims of the metaphysician, music's meaning is inextricably and irreducibly perceptual: inseparable from its sonorous presences, and undetachable from its sensible surfaces. To be understood fully, things must be perceived fully, and musical understanding cannot be separated from musical perception. Analysis and abstraction can never supplant the rich, embodied meanings that inhere in the perceived musical surface."
>
> (Bowman, 1998:263)

An example of how phenomenology deals and conceives music's meaning is given by Clifton's (1983) 'applied phenomenology'. Music, writes Clifton (1983:1), is "the actualisation of the possibility of any sound whatever to present to some human being a meaning which he experiences with his body – that is to say, with his mind, his feelings, his senses, his will, and his metabolism." Music is, according to him, a 'bodily phenomenon', i.e., it should be understood through the way its sounds are experienced through the body, the way they are lived through the movements and gestures that enact them, and its meaning constructed by an 'embodied mind'. Clifton (1983:41) writes:

> "Perception, is a two-way street. We are not passive receivers of uninterpreted sense data, nor are we the cause of an object's properties. Thus, while it is true that a sonata by Mozart exists independently of me, it has significance *for* me to the extent that I perceive it adequately."
>
> (Clifton, 1983:41)

Music is, thus, a bodily experience, a lived and performed experience that lies at the interface between the sound of music and the perceiver's full experience of them. The four

---

[26] Gibson states that the 'invariant structure' is whatever optical pattern persists, despite the changes of perspective; the 'changes' are what Gibson calls the 'transformations' (see here below, item 1.7 'The Ecological Contribution'). Gibson (1966) refers to the French folk wisdom - *La plus ça change, la plus c'est la même chose* - to characterise the meaning of the invariant and the transformation, for it describes how, by actively looking, the persistent features remain. As Gibson (1966) stresses: "something becomes different but it is not converted into something else" (quoted in: Reed & Jones, 1982:178). See further discussion in item 1.7 "The Ecological Contribution".

fundamental constituents of a meaningful musical experience of Clifton's applied phenomenology are time, space, play, and feeling, instead of the conventional categories of music analysis. Traditional music analysis, for instance, considers 'pitch' and 'interval' as essential to music. Clifton notes that pitch, is not essential music because there are many music repertoires where pitch has a minor role. Clifton (1983) asks:

> "Do we really hear *col legno* as something simply attached to the primary substance of pitch?
> (...) If a French horn prolongs an open E, and then quickly mutes it, is it the same E? Logically,
> yes; but in terms of musical behaviour, I think not."
>
> (Clifton, 1983:6)

Of course, Clifton admits, pitch may function as a medium, something through which one encounters music, but it is an essential feature of music. According to him, distance between pitches – 'intervals' - are musically meaningful only as a gesture:

> "When I hear an 'octave', (...) I hear a form of musical space, experienced either as thickening
> of a single line...or as the stratification of an otherwise undifferentiated space..."
>
> (Clifton, 1983:22)

'Harmony', though experiential in its capacity of attracting consonant tones and repelling dissonant ones, is first of all a conventional medium and therefore not an essential one. 'Tonality' is also, according to him, not one of the essential 'musical strata':

> "If I did not already have a prior understanding of motion in time, of 'toward' and 'away from',
> rest and tension, beginning and ending, anticipation and fulfilment, how would I ever recognise
> in myself the power to appropriate precisely these tones, through the medium of which a certain
> kind of motion, etc, called 'tonality' is particularised?"
>
> (Clifton, 1983:33)

Thus, he concludes, tonality is a 'corporeal acquisition', the achievement of a body that "knows how to behave in a tonal situation" (Clifton, 1983:34). The sense movement within tonality is directly related with body feeling and body movement. Without that, gestures become mechanical and music an external object. According to Clifton (1983:35), "if descriptions of tonality are so full of feeling-words it happens because in tonality 'I' is at the centre of it; and the tones which 'I' perceive as swirling around this centre have significance only because of the presence of that 'I' which sustains them." Clifton's point

of view is that we should consider music as a 'fundamentally bodily phenomenon', and not a kind of phenomenon in which music just refers to something bodily. "Tonality is a movement of my body", says Clifton (1983:34). That's why the symbolic accounts of music are also problematic for him. The idea that music's significance consists in a relationship between music and 'something that is not music', but to what music resembles very closely, is not an accurate one. Clifton believes that there has been through metaphysics influence an unnecessary demarcation between intramusical and extramusical attributes of musical meaning. Feelings, he states, are qualities that belong to music. 'Tonality' and 'tenderness' are both equally attributes of music experiential field:

> "If I become tender and dignified, it is because the music is tender and dignified; if I am tonal, the music exhibits the pull and tension of tonality. In the presence of music, I qualify my own ontology: I am tender and dignified; I am tonal."

> (Clifton, 1983:282)

So, according to Clifton's applied phenomenology, music is not mere auditory stimulus, its perception is fundamentally synaesthetic. The body, says Clifton (1983:68), is a "synergetic system which responds to a musical situation addressed to tactile and visual, as well as auditory, functions. (...) It is my body which produces its effects on tone quality, and it is only because of my body that there can be any talk of quality at all." A similar point of view has been articulated by Lochhead (1995):

> "Music perception is not the mechanistic process of registering the properties of an independently world or...of interpreting the data of sensory input, but rather it is a *bodily enaction of meaning*... I meet the sounds with my body and through it enact the melody as a felt significance."
>
> (Lochhead, 1995:36)

Thus, it seems that, in a phenomenological account, for example synaesthesia is not only fundamental to musical experience but, also, it is the source of all musical knowledge. Music is thus an 'event', a bodily experiential event and not just an objective and formal thing. As music is experienced through the body, music, engenders a kind of synaesthetic identification, engaging the body in a quite specific mode of articulation, which according

to Clifton (1983:144), could be considered the auditory counterpart of visual pattern: the 'gesture'. A parallel could be here drawn with Bierwisch's (1979) notion of "gesture" related to the meaning of musical signs. According to Bierwisch (1979:55) "gesture is directly and in and of itself a temporarily structured pattern." Movement in music (for instance, movement of pitch) represents thus, and according to Bierwisch (1979:56), "movement in emotions through something like an abstract 'space in motion' which can be grasped through its synaesthetic relatedness to the physical field of movement, to its characteristics of gravitational force, distance and height." In the lived musical experience of Clifton's phenomenology, music is not a physically objective thing but an 'event' that engages one's entire body in a "gestural activity" (Clifton, 1983:143). Thus, we may concur with Clifton that even the feelings perceived and expressed in music are not things that music refers to, or stimulates, or even that music contains. Feelings and emotions are essential to music's very being, a fundamental and inextricable part of what music is. Music, writes, Clifton (1983:288) "is then indistinguishable from the attitude which it induces." Bowman writes, at this purpose, that:

> "Music's fundamental sense of spatiality is a function of one's bodily experience of a phenomenal field. It is active, not contemplative; synaesthetic, not merely aural. (...) Musical experience always takes place in a multidimensional field of action and motion. (...) Thus, tonality is a sensuous rather than a syntactical affair. Its centre is not pitch, but one's body: a body that moves in space and whose fund of spatio-temporal experience ('toward' or 'away from'; inside and outside; centre or periphery) is what enables one to experience tonally in the first place."
>
> (Bowman, 1998:274)

To resume, the importance of Clifton's phenomenology can be seen as a contribution to unify the musical experience and to neutralise the mutual exclusivity of the subject/object and of other dualistic categories (mind/body, intramusical/extramusical, etc.) of the idealistic legacy. Clifton's account allows for the possibility to consider music as a field of human knowledge where music's significance is entirely based in the human bodily experience. Nevertheless, the idea of an isolated 'I' supporting all the experience and meaning of music might create the possibility of equally valid different experiences allowing the fact that music might never be seriously implied in the meanings it engenders.

1.7   THE ECOLOGICAL CONTRIBUTION

Gibson's (1966, 1979) 'ecological approach' is based on the idea that the environment is rich in information about its structure and dynamics and his main theoretical proposal is that information is by that means directly available to the perceiver. That is, perception constitutes some kind of awareness of what is specified in the environment without recourse to a series of constructive stages. Gibson (1960) suggests that in order to perceive a stable environment, one must be able to detect constant, invariant elements in the visual and acoustical array. As Shove and Repp (1995:62) explain, "perception is, thus, based not upon discrete sensations, but upon the 'pickup' of 'stimulus information', or simply, information. ...(This) information in light (or sound) is not concrete like the objects and events it specifies. Rather it is abstract and formless, consisting of relational invariants defined over transformations, invariants of a higher, mathematical order." In Gibson's (1973) own words:

> "Continuous optical transformations can yield two kinds of perception at the same time, one of change and one of non-change. (...) Object perception does not depend on form-perception but on invariant detection. And these invariants are 'formless', that is to say, they are not themselves forms."

> (Gibson, 1973: 43)

These 'invariants' consist, thus, of certain ratios and proportions that can be directly detected (that is, without further cognition) by the perceiver. For instance, all the elements one needs to specify the cause of, say, an acoustic event, should be available in the 'wave train' (the complex frequency array specifies the vibratory event and the sequence of transients provide information pertaining to the temporal profile of the event). According to Clarke (2001:219), "the ecological approach asserts that the source of any stimulus information is specified as it arrives at the sensory system of a perceiver: its spatial location, material, separation from other sources, and so on." Though Gibson argued for the direct availability of information in all senses, his own empirical work and theory focused mainly on vision. Gibson's theoretical point of departure is, in what concerns vision, that radiant light from sources like the sun is reflected from the surfaces of the environment and is structured by them. Let's take an example, quoted by Repp & Shove (1995:62), to

explain further what has just been meant: "A Child cries, thus transforming the contours of his or her face, yet one continues to see the same face." According to Gibson, the optic array (the flowing pattern of light) must contain information 'specific' to both the changing and the persisting features of the face. Thus, and according to Gibson, informational structure specific to the unchanging properties of the environment is invariant under transformations of the optic array and is thus revealed by the transformations. Source specification of a sound can surely be investigated in the acoustic stimulus information in a similar way to which the structure of any 'ambient optic array' is supposed to be uniquely specific to its source, and therefore informative about that source. According to Clarke (2001):

> "The ecological approach asserts that the source of a sound is specified in the stimulus information as it arrives at the sensory system of a perceiver: its spatial location, material, mode of excitation, separation from other sources, and so on. [...] Since sounds in the everyday world specify (among other things) the motional characteristics of their sources, it is inevitable that musical sounds will also specify movements and gestures, both the real movements and gestures of their physical production (...) and also the fictional movements and gestures of the virtual environment (*cf.* Windsor, 2000) which they conjure up."
>
> (Clarke, 2001: 219-220)

Therefore, in Gibsonian terms, it may be concluded that all movements (including those gestures symptomatic of human emotion) are also embraced by the specification process of both everyday sounds and musical sounds. In fact, the work of Fónagy (1967), Fowler (1986), and Sundberg (1987) seems to suggest that sound can specify movement and even movement symptomatic of human emotional states. That is exactly what Bregman (1990) claims, when he writes about the foundation of experience of sequential form in music as much as in real life:

> "Transformations in loudness, timbre, and other acoustic properties may allow the listener to conclude that the maker of a sound is drawing nearer, becoming weaker or more aggressive, or changing in other ways. However, in order to justify such conclusions, it must be known that the sounds that bear these acoustic relations to one another are derived from the same source."
>
> (Bregman, 1990:469)

In essence, the 'ecological approach' recognises that light (or sound) is structured and therefore carries information, which is meaningful in respect to its source. That is why source specification can be regarded, as Clarke (2001:218) remarks, "as an aspect of meaning (perceptual meaning), and is as relevant to the perception of the cultural environment as it is to the natural environment". In this regard, Gibson (1966) is right to claim that there is no sharp discontinuity between nature and culture:

> "Symbols are taken to be profoundly different from things. But let us be clear about this. There have to be modes of stimulation, or ways of conveying information, for any individual to perceive anything, however abstract. He must be sensitive to stimuli no matter how universal or fine-spun the thing he apprehends. No symbol exists except as it is realised in sound, projected light, mechanical contact, or the like. All knowledge rests on sensitivity."

> (Gibson, 1966: 26)

As Clarke (2001:219) points out, "a great deal stands to be gained by doing so, however, and particularly, in bringing some kind of continuity and unity to the way in which we understand the perception of 'basic' features and the cultural meanings of music."

However, the perceiver's capacity to detect that information is not, according to Gibson, innate: "The perceptual capacities of the new born, animal or human, for getting information becomes a matter of investigation" (Gibson, 1966a: 267). Perception is, thus, a matter of learning and investigation. And this investigation is achieved by a 'seeking-out' with movements of the eyes, head and often the whole body, and responding to the transformations of the environment. According to Gibson (1958):

> "Perception is active not passive. It is exploratory, not merely receptive ... Exploratory movements of the eyes, and even locomotor exploration in surroundings may all be thought of as a search for more information."
> (Gibson, 1958: 43)

Gibson coined the term "ecological level" to name a level of perception which preserves as many features of the individual's natural behaviour as possible. This is the essence of what according to Gibson's theory has been called the 'ecological validity' of an experiment. The structuring medium, be it light or sound, does not so much carry as contain information and, thus, it provides a continuous flow of stimulus energy. In this regard Shove & Repp

(1995:62) comment that, "under favourable conditions this information is sufficiently available to the perceptual systems, thus enabling the perceiver to tailor his or her actions to the environment." This is the starting point for Gibson's theory. It emphasises the idea that with structured information in the ambient optic (acoustic) array the perceiver and the perceived are complementary. The perceiver detects the structured information as an 'immediate acquaintance' (Gibson, 1960), an awareness of what is specified in a whole perceptual system. To perceive is to be 'resonant' with the perceived. The perceiver, in order to pick-up the structured information, has to 'tune' into it. The human perceiver is a self-tuning system, and the adjustment of the perceiver's sensitivity or attunement to the object – affects what the object may mean. 'Attunement' is dependent on experience, which is achieved by the 'education of attention'. In essence, to educate one's attention is to refine the ability to discriminate through practice (Gibson & Gibson, 1955). That explains the different awareness and the different meanings that a same percept may offer or 'afford' to the same or to different perceivers.[27] In other words, the perceiver's sensitivity or attunement to the object affects what the object may mean. This explains why the same object may afford different properties to different people. Meaning of an object is thus directly connected with the perceiver's level of awareness.

That is the reason why, as Davidson (1991:37) comments, "the root of a culturally developed environment is the natural environment, which is modified to change what it affords, but this is not artificial and distinct from the natural environment: it is the same environment." But according to Gibson (1966), there are also misleading distinctions that should be avoid:

> "The cultural environment…is often divided into two parts, 'material' and non-material' culture. This is a seriously misleading distinction, for it seems to imply that language, art, music, law and religion are immaterial, insubstantial, or intangible, whereas tools, shelters, clothes, vehicles, and books are not."
>
> (Gibson, 1966:26)

---

[27] According to Gibson's terminology, the "affordances of things are what they furnish, for good or ill, that is, what they afford the observer" (Gibson, 1979:127). Cutting (1982) views this concept as an ideological masterstroke, for, in effect, to perceive an affordance is to perceive meaning (quoted in Davidson, 1991:35).

But, Gibson's materialism comes always associated with a no-less important anti-dualism:

> "An affordance points two ways, to the environment and to the observer. (...) But it does not in the least imply separate realms of consciousness and matter, a psychological dualism. It says only that the information to specify the utilities of the environment is accompanied by information to specify the observer himself, his body, his legs, hands and mouth. This is only to re-emphasise that exteroception is accompanied by proprioception – that to perceive the world is to co-perceive oneself. This is wholly inconsistent with dualism in any form, either mind-matter dualism or mind-body dualism. The awareness of the world and of one's complementary relations to the world are not separable."
>
> (Gibson, 1979: 141)

Thus, awareness of the world and awareness of oneself are complementary and inseparable things. Both are material and substantial things. Both are involved in the same process of meaning affordance. But, an affordance does not only offer up meaning, it also provides a motive for action. And, this action is not directly caused by the meaning perceived in the environment's events. Action is a complementary awareness that the perceiver develops of himself, as he becomes aware of the world. It is possible to find as a part of the perceptual experience two kinds of potential action: the 'exploratory action' controlled by a search of the affordances of the environment and the 'performatory action' as a response to those affordances and controlled by the perception of them. The relevance of the affordance concept can be seen in different approaches within several research fields. For example, Zebrowitz (1990) applied this concept to the study of emotion in social perception research field:

> "...emotion expressions may be viewed as specifying social affordances such as 'approach me' or 'avoid me' or 'help me' rather than simply as 'happy' or 'angry' or 'afraid'. Indeed, emotions are associated with specific states of action readiness... (and) a fearful or angry face not only signals that the environment affords danger but also facilitates appropriate adaptive actions by perceiver..."
>
> (Zebrowitz, 1990:182)

Zebrowitz (1990) concluded that the concept of affordance is able to provide not only an accurate account of the perception of an expression, but also of the resulting perceiver's action. Though Gibson's perception theory has mainly focused on the role of vision, his ecological approach, as Davidson (1991:38) points out, "stresses that all senses are capable

of detecting invariants and transformations, and therefore that all are involved in consequent motor adjustment." But, if Gibson has developed these ideas primarily in relation to vision, other authors in an attempt to apply the same ecological principles to hearing contributed to a further development of the ecological approach.[28] Bregman (1990), for instance, realised that the examples of auditory perception based on Gestalt principles are not isolated phenomena but are part of a larger scheme. Bregman coined the term "auditory scene analysis" to explain the mechanisms of aggregation and segregation that reveal acoustic information about the events that take place in the surrounding environment:

> "Sound is created when things of various types happen. The wind blows, an animal scurries trough a clearing, the fire burns, a person calls. Acoustic information, therefore, tells us about physical 'happenings'. Many happenings go on at the same time in the world, each one a distinct event. If we are to react to them as distinct, there has to be a level of mental description in which there are separate representations of the individual ones."
>
> (Bregman, 1990: 10)

A musical example will illustrate what has been just exposed. According to gestalt principles, to hear a melody when played by an instrument would mean that all notes have a similar timbre (principle of similarity), all notes fall within a reasonably narrow pitch range (principle of proximity) and all notes bring a sense of continuity and unity (principle of good continuation). A scene-analytical explanation would add that a listener has used cues such as the complex frequency array and the sequence of transients, beyond those of similarity, proximity and good continuation, to segregate out a sound source (for instance, a human voice) from the complex auditory environment. The ecological explanation would also add that, "to hear a melody is to hear, among other things, a musician move relative to the surface of an instrument (and that includes the vocal apparatus)" (Shove & Repp, 1995: 61).

---

[28] Recently, and complementing Gibson's work in 'ecological optics', an important number of musicians and music psychologists have approached the study of movement and that of the perception of movement in music performance. They have made significant contributions to what is called nowadays the 'ecological acoustics'. The reference to some of these major contributions will be focused in detail in the second chapter of this thesis. See Davidson (1991) for a detailed exposition of the argument for an ecological approach, the main proposals of Gibson's theory and the empirical evidence in ecological social perception.

So, the importance of the ecological approach is to favour the links between the person and the environment, and the organisation and structure provided in the external stimulus information. It offers a sensitive and rich framework for empirical research specially if one thinks of the possibilities of the ecological account in what concerns the expressive intentions manifested in body movements. Reed's (1988:313) statement highlighting the importance of Gibson's theory, brings back the central argument of this chapter, i.e., the importance of the elaboration of a framework for empirical investigation whose conceptual tools successfully avoid and overcome the irreconcilable dualism of the categories of the metaphysical legacy (mind/body, subjectivity/objectivity, universal/particular, culture/nature, inner life/outer world, intellect/feeling, reason/emotion, intra-musical/extra-musical meaning, etc):

> "James Gibson's most enduring achievement was to show that the study of perception could be both scientific and realistic. Ever since the scientific revolution, the study of perception has been crippled by the belief that we cannot see the meaningful things surrounding us, that what we see are mere images of the world or, worse, merely fleeting sensations that are caused to appear in our consciousness by the ultimately unknowable machinations of the outside world... From the beginning, Gibson challenged the basic assumptions of this unfortunate heritage. It is typical of Gibson's working style that he was the first to understand and push beyond the limits of his own critique... Gibson felt he was merely showing the way to a new range of problems and issues, not presenting a settled and complete theory."
>
> (Reed, 1988:313)

In order to compare and examine further the possibilities and practical consequences of some of these main contributions to the study of musical meaning, and specifically to the study of emotional meaning in music/performance, it is necessary to turn to more practical aspects of the relation between theoretical accounts and empirical investigations. In this sense, Chapter 2 presents a review of some of the most relevant theoretical-practical frameworks and models elaborated for the empirical investigation in music/performance.

1.8   SUMMARY

In this chapter, an assessment of how the metaphysical legacy of idealism is still present in most of the contemporary theoretical accounts of music's nature and value has been given, and also how it still 'contaminates' some of the conceptual tools that make part of the empirical approach to the meaning of music. It has been also an aim of this chapter to examine some of these theoretical contributions to facilitate an understanding of musical meaning, nature and value, and to recover and refresh some of their ideas and concepts in order to explore them as conceptual tools for the empirical investigations to be undertaken in this thesis. Concepts like expression, meaning, intention, structure, reference, perception, have been used as conceptual tools in the different theoretical accounts to pursue and explain the multifaceted nature of the musical experience. These concepts have been examined within the theoretical context of their appearance, and in terms of the most up-to-date knowledge on the empirical research in music performance. For the sake of clarity, they have been presented and named as theoretical contributions to the comprehension of music. They have been investigated as follows: the *iconic contribution*, the *autonomous contribution*,   the   *structuralist   contribution*,   the   *symbolic   contribution*,   the *phenomenological contribution*, and the *ecological contribution*.

Within *the iconic contribution*, and despite its vagueness, the idea of a 'similarity', or a 'homology' between the 'dynamics of music' and the 'dynamic qualities of feeling' has become an important topic for theorists and researchers all over the world. Despite the limitations of its historical appearance (empirical music research had just started) this concept can be seen as a direct precursor of the concept of 'vitality affects' that were presented by Stern (1985). Its importance to this thesis can be seen at the level of understanding how musical structure might be able to start a process of emotional meaning by generating a kind of 'dynamic' expression able to evoke the dynamic content of human feelings and affect.

The *autonomic contribution* (mostly based on historically dated research - Dewey's (1984) "conflict theory" of emotions, and on Rapoport's (1950) psychological findings - conceives music's expressiveness as being primarily directed to the way listeners, emotionally and/or rationally, make sense of the interplay of the musical structures and of the syntactical relationships among musical events. These concerns brought the idea of music expressing emotion within the objective relationships and structures of music, and

created the opportunity of bringing emotion objectively into the realm of music and of music theory.

The *structuralist contribution* assumes that meaning and signification are contained within a particular piece's structural relations. The musical object is viewed as an empirical entity to be probed by the subject in such a way as to reveal its underlying systems or coherence upon which its perception as 'meaningful' rests. Unable to relate accurately the musical meaning with the extrinsic world of objects and events, structuralism leaves us, nevertheless, with the value of a new concept, the structure, which is able to defend the integrity of music as a distinct and particular form of human expression and communication. Its importance in the scope of this thesis is also connected to the fact that most late empirical research has been investigating how music performance relates expressively to musical structure.

The *symbolic contribution* aid to clarify the way in which the internal states, somatically experienced interrelate and intersect with the sounds of music, and 'become source of externally uttered movements'. This is one of the most exciting sources of contemporary musical performance research. Player's movements affect musical structures in an effective way and the motor patterns implicated in different kinds of music change according to the particularities of each music because they are the result of the interaction between the morphology of the instrument and the player's sensorimotor capacities (Baily, 1977). This is especially interesting, since the central subject of the research of this thesis is the human voice, and how it articulates within singing the investment of emotional meaning.

The importance of the *phenomenological contribution* can be seen as a contribution to unify the musical experience and to neutralise the mutual exclusivity of the subject/object of the rational proposition. The phenomenological account allows for the possibility to consider music as a field of human action where music's nature and value is entirely based in the human bodily experience.

The importance of the *ecological contribution*, historically a much more recent theory, is to emphasise the study of encounters with one's environment and the idea that the environment is rich in information about its own structure and dynamics. Such an approach means developing experiments which preserve as many features of the individual's natural

behaviour as possible. This is essence of what Gibson has called the 'ecological validity' of an experiment, and this has been an aim in the empirical research of this thesis, especially in the scope of Experiment-IV.

Despite the differences in focus within the contributions above, a common idea comes across the different theoretical accounts. This idea is connected with the perception of motion in music and with the fact that music seems to be able to somehow express and mirror the dynamics of human inner life or inner movement. Of course, the idea of the existence of a certain iconicity or homology between the music motion and the human movement, be it movements symptomatic of emotions and feelings, or any other form of human expressive movement, is not seen in the same way within the different theories. The differences found in this interface between music and motion, reflect thus the different modes of musical awareness among the several theoretical contributions exposed.

CHAPTER 2

2. THEORETICAL-PRACTICAL BACKGROUND

REVIEW OF SOME KEY THEORETICAL FRAMEWORKS ELABORATED FOR EMPIRICAL APPROACHES TO THE STUDY OF THE EXPRESSION OF EMOTION IN MUSIC PERFORMANCE (SINGING).

2.1 INTRODUCTION: MUSIC (SINGING) AND EMOTION.

According to Small (1987) Western singing might be seen as a complex ritual where singers and audience exchange meanings within a cultural, aesthetic and social space/time. The space/time of singing is a 'ritual happening' where cultural expectations and social codes rule the communication between singer and audience, but also where cross-cultural elements and a infinite set of body movements and affect expressions are performed. Thus, performance actions, gestures and sounds may be seen as a part of a process where social, cultural and aesthetic aspects of communication and interaction with an audience are brought up through the art of singing.

From a historical point of view, it seems also that in Western Art Music performances, the audience seems confined to absolute social and cultural conformity. Looking at historical context and observing performers nowadays who sing for audiences, it is easy to recognise that the expectations placed upon them are accorded to the cultural, aesthetic and social demands of our century, and these are very different to the expectations of, for instance, Monteverdi's century. It seems that, from a contemporary point of view, the weight of a good music performance is more and more placed on the technical skills and on the musical expertise of the performer (Newham, 1998). Expressiveness, another apparently important musical factor in music performance, seems to be receiving little attention from contemporary performers, audiences and, even, from performance literature (Hemsley, 1998). But perhaps Caccini, one of the major exponents of the dramatic style change that appeared in Italy in the XVI century, might contribute to re-capture the neglected importance of this other major role in musical performance: that of expression. In his introduction to Le Nuove Musiche, Caccini (1614) noted, that music must honour the natural contours of the vocal line, which in turn must serve the affective expression and the communication of emotion. Of course, these imply natural human phenomena grounded in physiology rather than sociology.

It is generally accepted that in the history of the Art of Western Singing that vocal specialisation appeared rather late, and as a consequence of the loss of the Bel Canto tradition and ideals in XIX century. Bel Canto was regarded as the ability to communicate the human emotions by singing precisely notated musical phrases with a wide range of qualities based on the exclamatory vowels – 'ah' for pleasure, 'ee' for disgust and hatred, 'oo' for fear and horror - as a means of interpretation, extending across a wide range of several octaves (Manén, 1974). It seems, therefore, that in this tradition physiological considerations were being interfaced with socio-cultural premises.

The **Bel Canto** singer was supposed to be able to sing a vast array of qualities covering all different kinds of the traditional vocal compartments. The voice was meant to have not three, four or five sonorities but in fact twenty or thirty. The voice had to be moulded in an infinite degree, passing through all the colours of a 'sound prism' (Hahn, 1920). Whist Bel Canto singers were intent on mastering the art of extending the different possible emotional qualities and imaginative characteristics on each note, it seems that nowadays, singers seek to specialise in the perfection of a single operatic quality of voice and try to extend the range associated with it. According to Newham (1998):

> "Later, singers began gradually to have to cope with operatic roles which required more specialised qualities of sound, and this led to a process of training the voice which was more influenced by the technical demands of the music and less connected to the primal and fundamental role of the voice as the expression of emotion."
>
> (Newham, 1998: XX)

So, different socio-cultural demands were placing a very specific form of constraint on the physical/cognitive approach to voice production. But, if today the expression of emotion seems to have lost its leading role in the performance of singing in Western Art Music Canon, emotion has not completely vanished from singers and others musicians' performances and vocabulary. One of the most discussed issues among singers and their audiences (Draffan, 2000) is certainly the issue of "if" and "how" emotion is being communicated within the process of singing. It is also one of the most important issues in terms of a singer's self-evaluation and self-identification, and on the singer's sense of a

'truthful' performance, i.e., a performance where s/he has feelings of oneness and wholeness with the public. Corroborating this idea is a long series of empirical work on the subject of the impact of self-concept, confidence, emotion and social environment on the singer's voice and ability to perform by Draffan (2000). Here are some indicative examples of data she generated from interviewing professional singers about their work and feelings towards it:

'When I have done something that really feels like the sort of performance I believe in where you know that you're a channel and that somehow you haven't even been singing because its gone through you; these moments are the handful of almost meetings with God.... That utter feeling of nothing being in the way of your heart and the rest of the world, which is a complete high you don't get from very many things, and I think I've only ever had it through listening to, or performing music. '

'You discover yourself through singing, what you are capable of, what your body can do and what you can do with your imagination. '

'If I can express myself in my life utterly truthfully so that everything I do in my life is completely, honestly and authentically me then there wouldn't be any problems with anything I chose to do and I would sing in a very different way. '

'Singing is very emotional. When I was finding out about it 3 years ago, if it was feeling closed I'd open it up and I'd just start crying. It was 'nice' crying and I'd keep singing and I'd feel better. It was like just getting to a place where you feel you were being honest and real. '

'Most people in life, who are working in 'normal' jobs, haven't got to invest the emotional level into their career ladder. '

'It [singing] must be the biggest trigger for any amount of emotional rubbish that might ever come out of me. '

'If what you are trying to do with your music is really heartfelt, then every time you put your voice out there, with all the emotion attached to it, you're just up there to be shot down as a whole person. '

'It [the voice] *can express a lot of your frustration, your happiness or sadness. It is directly connected to emotional states and problems.'*

'It is amazing the amount of people who have been going through crises and have not been able to sing. Not having the will to sing and not being able to physically. And people who have lost their voices through this. So it has to be linked to the emotions.'

(Draffan, 2000:33)

Thus, the communication of emotion between performer and audience seems to be still, and despite all contrary tendencies, at the very heart of the performance of singing. It is surely the case of all other performance experience too (see Persson, 1995), but it is especially with singers that most of the time the issue of the expression of emotion appears as an important, or perhaps, the most important content of a good performance, since it is a direct communication of the music and its intention with the performer's body/self.

The acknowledgement of the strong relationship between music and emotion was recognised from as far back as the ancient Greeks' time, [28] and had a further development in the discussions and policies of the Music of the Middle Ages. For instance, the affective qualities were set for Gregorian modes for religious expression (Epstein, 1993). Late Baroque and Early Classical music clearly aimed at expressing and communicating emotions. The 'Affektenlehre', doctrine of the affections (Mattheson, 1739), tried to establish, during the Seventeenth and Eighteenth Centuries, musical effects that could be used to express specific emotions – joy, sorrow, passion, and languor are just some examples of these. In the late Eighteenth Century 'affect' was regarded as a rationalised emotional state in contrast to Nineteenth Century were emotion was a personal and spontaneous expression (Cook and Dibben, 2001). Small (1998) has accounted for the

---

[28] "Musical training is a more potent instrument than any other, because rhythm and harmony find their way into the inward places of the soul (...) He who received this true education of the inner being will most shrewdly perceive omissions and faults in art and...while he praises and rejoices over and receives into his soul the good...he will justly hate the bad, now in the days of his youth even before he is able to know the reason why (Plato, *The Republic*, Book III, in Jowet, 1952: 289). And, also, when referring to musical rhythms, Plato declares that 'paideia' should take in careful account "what rhythms are expressive of meanness, or insolence, or fury, or unworthiness, and what are to be reserved for the expression of opposite feelings" (Plato, *The Republic* III, in Jowet, 1952:399).

abstraction and rationalisation of embodied meaning in Western Classical Music as a result of the referred historical circumstances. He suggests:

> "[M]usical gestures were abstracted from physical movement so that the listeners no longer moved their bodies [no dance movement responses] but sat and watched and listened, and (...) the musical gestures represented not an emotional state itself nor a temperament but the type of physical gesture, both bodily and vocal, with which the emotional state or the temperament was associated. The musical gesture represented metaphorically the physical gesture that the audience recognised as belonging to that state. It thus had to be constructed at one remove, and the masters of that first brilliant explosion of the new art form worked through conscious striving, exchange of ideas, polemics and a good deal of trial and error, to perfect the representation."
>
> (Small, 1998:148)

The Nineteenth Century brought, also, the development of programme music to expand musical expressiveness not only to more complex feelings and emotions but, also, to events, ideas and philosophies. However, it was the aesthetics of Renaissance the very core of the view that attributed emotional effects to music. Vincenzo Galile, a member of the Camerata of Florence, wrote:

> "Music exists primarily to express the passions with great effectiveness and to communicate these passions with equal force to the minds of mortals for their benefit and advantage."
>
> (Galile, 1581: 306)

But how or by what means was music supposed to express passions, to convey or arouse emotion, and to communicate it to an audience? The answer to this question urged the need, over the last few decades, to elaborate a theoretical framework that could bring a scientific based explanation to the music-emotion interface. Of course, the theoretical accounts that try to enlighten the interface between music-emotion, diverge according to where the focus of the explanation is placed, i.e., the music features, the composer, the listener, the performer, the cognitive or arousalist aspect of the emotion conveyed, etc. The differences between some of the several theoretical accounts will be next reviewed and confronted.

2.2 REVIEW OF SOME THEORETICAL FRAMEWORKS ON THE STUDY OF THE EXPRESSION OF EMOTION IN MUSIC /PERFORMANCE (SINGING)

2.2.1 MUSIC AND THE EMOTION CONVEYED IN MUSIC/PERFORMANCE (SINGING)

A) AN 'ART OF CONTENT' MODEL

To the question how were the required emotional and ethical effects of music to be produced at the Renaissance, Beardsley (1966) answered:

> "...by an increase in musical resources: a richer harmonic language, mixtures of modes, modulations from one mode to another, the introduction of new instruments with wider tonal ranges, etc. But more importantly, through a subjection of music to its text. Since words are obviously the best means of arousing passions and conveying ideas, they must be given primacy in the song; the music must be made to follow the meaning of the words, to underline and intensify the moods they express. The emotional resources of melody, harmony, and rhythm exist to increase the affective language of the poet."

> (Beardsley, 1966:131)

By subjecting music to its text, and by increasing the affective language of the poet through the emotional resources of music (melody, harmony and rhythm are supposed to underline and intensify the meaning of the text) the Renaissance artist has approached music associated with verbal text to the so-called 'art of content'. Like literature, poetry, drama and representational painting, music when associated with verbal text is to be explained by the fact that it represents the *universal* by means of the re-presentation of a particular experience. And, like other 'arts of content', it seems to conform to the theory of the *narration-imagination-identification,* one of the three theories identified by Noy (1993), that might constitute a methodological framework for empirical research in order to explain the emotional effect of music. These theories are known as the theory of *narration-imagination-identification,* the theory of *isomorphism* and the theory of *ego mastery.* The theory of the *narration-imagination-identification* represents the narrative route, predominantly used by the 'arts of content'. The other two, the theory of *isomorphism* and the theory of *ego mastery* represent, respectively, the direct route and the indirect route,

predominantly used by the 'arts of form'. [29] According to the principle of the *narrative route* or route of *narration-imagination-identification*, the artist searches to succeed in evoking his/her audience's emotion, luring the audience to identify with the experiences and emotions appropriate to the narrative content by arousing imagination:

> "One becomes involved in the story as if it were one's own experience; by identifying with one or several of the heroes and their deeds one experiences the emotions appropriate to the situation... This route is clearly involved in all music associated with verbal text, such as the lied, opera, ballet, and program music in general."

> (Noy, 1993:128)

According to this idea, singing in its operatic, song, and oratorio specialisation can be very much seen as an 'art of content' because it ensures the audience's identification by means of a content associated to a verbal text and/or to a clear musical narrative. The singer communicates the musical/verbal narrative in such a way that activates the audience imagination. The public ideally becomes involved with the story as if it were their own experience and by identifying with the character/s and their actions experience the emotions appropriate to the situation. Faulkner (1999) corroborates this idea by writing that:

> "There is a philosophy of music which argues that all Western music more or less since Monteverdi's Orfeo has been representative. If listeners and performers experience music as *stile rappresentativo* it would seem likely that this narrative, with characters, images and events might be the source of expressive intention and of its perception. Nowhere is the role of the narrative more explicit than in singing."

> (Faulkner, 1999: 7)

Elliot (1966-7) also seems to corroborate Noy's explanation, but adds another nuance when talking about the way emotion is evoked in the audience. He states that any art that uses the narrative route to convey emotion will be able to evoke the experience of an emotion imaginatively from within and from without. According to Budd's (1985:126) explanation, to experience the poem from *within* is "to place oneself, in imagination, in the situation of

---

[29] According to Noy (1993:125): "These three theories represent three different principles for explaining how art may convey emotions, or three different routes by which the artist may succeed in evoking his or her audience's emotions: the *narrative route*, the *direct route* and the *indirect route*".

the poem's 'speaker' and to experience the expression and the emotion expressed from that position." To experience the poem from *without* "one imagines someone giving expression to his emotion in the poem but one does not identify with him imaginatively." Elliot makes also a distinction between an emotion 'really' experienced and an emotion 'imaginatively' experienced by introducing the use of the words "primordial' and 'non-primordial'. This distinction between *primordial* and *non-primordial* seems to be of some relevance to the interface between 'real' emotion and 'performed' emotion. According to Elliot, we may say that every time a song or a poem expresses an emotion we may experience it from within or from without, but not in the primordial way. Thus, every time 'a performer' experiences an emotion when performing a song or a poem from within or from without we may say that 's/he' is experiencing an emotion in the non-primordial way.[30]

B) AN 'ART OF FORM' MODEL

So, if singing is seen, like poetry or drama, as an 'art of content', it should be noticed that part of the creative and meaningful experience within the performance of singing is connected to the fact that singing deals with music. And music in its more abstract presentation, by means of its structure, can certainly be seen as an 'art of form'. Of course, form itself may have some direct effect on the receiver to arouse or induce an emotion, or an emotional state (Gabrielsson & Lindström, 2001). Differently from 'pure' music or a declaimed poem or text, singing connects poetry and verbal language with music, offering up a complementary musical and artistic field. This complementary musical field creates an opportunity to combine the *narrative route* with the two other *routes* (the *direct route* and the *indirect route*), in order to investigate how an emotion is perceived in music or, even, how emotional response is evoked in the audience by some music structural feature.

---

[30] Budd (1985) suggests that this experience (the experience of an emotion in the non-primordial manner) can be clarified by introducing the notion of 'make-believe'. According to him, "either one can imagine being uneasy or distressed at the realisation or the thought that one is in danger or one can experience unease or distress in imagining oneself to be in danger. The first of these imaginative exercises is imagining oneself to be afraid: the emotion is imagined but not necessarily experienced. The second is experiencing fear imaginatively: the hedonic tone of the emotion is experienced when the thought-content of the emotion is imagined as being true. A similar distinction can be drawn in the more complex case in which a work of art is used as a prop in a game of make-believe. Make-believedly one is afraid if it is make-believedly true that one is uneasy or distressed at the thought that one is in danger. One experiences fear make-believedly if one experiences uneasy and distressed as a result of the realisation that make-believedly one is in danger" (1985:128-132).

Singing seems to be at the intersection of an "art of content" and of an "art of form". In all three above referred styles, - opera, song and oratorio - singing involves verbal and/or dramatic narrative, as an "art of content", and music form and structure, as an "art of form". In that way, the singer has a privileged chance to combine the expressive means of these two groups of arts. It seems, thus, relevant that the singer should know which routes belong to each of these groups in order s/he can understand better the expression, the perception and the evocation of emotion in singing. Nevertheless, the ability for expression is, dependent not only on the expressive means the artist chooses, but on his/her creative ability to utilise even the most simple means to convey emotion while telling us a 'truthful' story. Shaffer (1992:25) argues similarly that some sort of narrative at least partly generated by the performer may determine the pattern of musical expression and mood over the events of the piece. In a wider sense it is acceptable to say that every art involves all three routes in conveying emotional meanings. The difference between one art and another can be seen in the relative proportion each route is used. In this sense, music, and particularly singing, for the exposed reasons, can be seen as a model for other arts.

According to Noy (1993), the capacity of the listener to respond to the music emotionally can be achieved by the narrative route, as already referred above, but can also be achieved "through the emotional impact of the listener, of the very building blocks of which this narrative is constructed (Noy, 1993:134)." It's the, so called, *direct route*. According to Noy's explanation, with the *direct route* the artist will try to have impact on the listener emotions, by using the following means:

> "Using messages that may activate *directly*, by means of their form, the emotions. (...)
> The group of direct means include all the musical elements that may allude to,
> represent, or transmit some emotional meaning, or act as stimuli for arousing or
> enhancing an emotional response in the listener. (...) This can be achieved by means
> that represent some meanings (via representational means) or by means that do not
> represent meanings (via non-representational means). (...) The first group includes all
> those means that affect the listener, owing to the specific meaning they are presenting;
> the second group includes all those means that affect the listener by activating directly,
> sometimes even accidentally, some of his emotions".
>
> (Noy, 1993:133-134)

Included in the first group (via **representational means**) are the two major theories that according to Kivy (1980) search to explain the expressiveness of music:

- The *contour theory*, which considers music to be able to provoke an emotional response on the listener owing to the "congruence of musical contour with the structure of expressive features and behaviour"

(Kivy, 1980:77)

- The *convention theory*, which explains the expressiveness of music "as a function, simply, of the customary association of certain musical features with certain emotive ones, quite apart from any structural analogy between them"

(Kivy, 1980:77)

These two theories may be seen not as opposites views of artistic expressiveness, but as complementary levels of general human communication. They might be seen as a two-level system that participates in the process of information conveyance in human communication. Noy (1993:135) argues that the *secondary level* "includes all symbols, signals, and signs conveying the meanings that had been assigned to them by agreed-upon social convention", while the *primary level* "includes all means expressing directly and concretely human emotions, experiences, and other subjective states of mind". On the secondary level "the objective knowledge is transmitted by means of language". On the primary level "the subjective experiences and feelings are expressed by means of all the accompanying gestures, inflation of voice, movements, and so on... Most of the experiences and emotions accompanying verbal communication of ideas are expressed by means of the tone, pitch, intensity, rhythm, resonance, timbre, duration, and inflection of the speaking voice, in short, by all we use to call *the music of speech.*"

Of course, the affinity of human speech and the expressiveness of music have long been observed. As Noy (1993) states:

"a considerable part of music's representational means conveying emotional meanings derive their meanings out of these "musical" components of speech. "Music uses a

variety of techniques for representing all kinds of meaningful non-verbal signs comprehensible to any user of language, such as imitating various well-recognised expressions of emotions, like laughter or weeping (the laughter of the orchestra accompanying the aria of Leporello in Mozart's Don Giovanni or the weeping of Dido from the Lament of Dido and Aeneas by Purcell), or by manipulating any of the expressive means presented above (i.e., tone, pitch, intensity). For example, Berlioz in the "Lacrymosa" of his Requiem, op. 5, presented the experience of weeping by imitating the non-rhythmical sighs characteristic of sobbing by means of disturbances in the line of regular rhythm."

(Noy, 1993:135)

In fact, enough empirical evidence (see, for instance, Scherer, 1986, 1995; Gabrielsson & Juslin, 1996; Juslin, 1998) seems to corroborate the idea that music/performance is heard as human utterance, i.e., sharing the same emotional profiles of vocal utterance, and the same physiological concomitants of emotional gestures. Thus, as Lavy (2001:54) argues, we hear musical sounds as human sounds and respond to them as though they are vocal utterances.

Included in the second group (via **non-representational means**) are all those theories that consider music (and art in general) to have "the power to arouse an emotional response in the listener in some direct way that is not dependent on any communication of meanings" (Noy, 1993:138). Noy considers the theory of *isomorphism* of Langer and Pratt the only current theory that attempts to explain comprehensively this phenomenon (see at this purpose the review of Langer and Pratt's accounts in Chapter 1, item 1.1 'The Iconic Contribution').

In order to explain the phenomenon of emotional responsiveness to music, Noy (1993) argues that "the communication of affect is based on genetically determined constitutional links. We are programmed in advance to respond to the perception of certain emotions with the arousal of our own emotions (Noy, 1993:140)." And suggests the term of *primary empathy* to "this kind of constitutionally determined human sensitivity and responsiveness to the manifestations of other's emotions" (Noy, 1993:140). To conclude, Noy (1993:141) concurs with the theory of isomorphism based on the assumption that:

"most of the signs indicating the activation of a given emotion can be communicated also by an alternative medium other than the original one; for example, the signs of mounting rage that can be communicated by bodily movements going faster and faster, increasing sounds, flashes of light, whirling red colours. (…) The novelty of the theory is in assuming that music and perhaps any art, can represent in its specific medium not only any of the particular changes (signs), but also the form of the entire pattern of changes."

(Noy, 1993:138)

With the *indirect route* approach, "music is treated as if it were only an array of acoustic stimuli that does not include, transmit or allude to any emotions but provokes the listener into some kind of organisational activity which may invoke in him some emotional response. The composer provokes the receiver's ego into that particular organising activity that, according to the artist's experience, may result in the emotional response the artist intends to arouse" (Noy, 1993:142). The best example of a theory using an indirect route, as an explanation of emotion responsiveness to music, is Meyer's theory (extensively referred to above). Noy (1993) argues that according to the indirect route approach, and in order to understand the emotional effect of a certain music, we have to look "into the listener's reaction to that music; not asking what the music is doing to the listener, but rather, what the listener is doing with the music." [31]

The theoretical framework just reviewed tries to enlighten different ways by which music is able to convey emotion. It is, though, not the only approach aiming to understand and explain the different modalities in which emotion and music interrelate in order to explain the emotional effect of music.

---

[31] According to him, this approach can not be presented as an alternative theory to explain the emotional impact of music (and art in general). It is, in Noy's (1993:142) words: "an additional explanation to the former two in advancing that music, in addition to telling a story and in addition to evoking directly an emotional response, may also provoke the listener into doing some active work."

2.2.2 MUSIC AND THE EMOTIONAL RESPONSE TO MUSIC/PERFORMANCE (SINGING)

A) A SEMIOLOGICAL MODEL

Nattiez's (1990) *music semiology* and Shepherd & Wicke's (1997) *music performative semiological model* are two main examples of theoretical accounts that consider an approach to the emotional response to music based on a semiological perspective. As both theories have already been extensively enough exposed in Chapter 1, a third theory offering up a general framework for empirical research on music and emotion based on a semiological theoretical background by Dowling & Harwood (1986) will be now briefly and comparatively considered. This theory offers a wide perspective on what in music might evoke an emotion in the listener. Dowling & Harwood (1986) draw on the work of Peirce (1931-37), and interrelate emotion to music at three different levels, that of 'index', 'icon' and 'symbol'. [32] According to them, music acting as 'index' works in a Pavlovian (1927) way, evoking emotions by provoking associations in a listener's mind between the music and an extra-musical emotional referent (it is the quoted example of Proust's (1934) violin sonata, when the recognition of a violin sonata evoked the memories of a love affair years before). Nevertheless, the *indexical level* does not need to be thought as just working with idiosyncratic associations, any sound, musical or non-musical, to use Peirce's terminology, can be the 'smoke' of an emotional 'fire', or, at least, the 'print' of it. This is, specially, true if one thinks it at the level of musical performance expression. In fact, not all the semiological models consider the indexical level as the level of idiosyncratic associations. Shepherd & Wicke's (1997) consider, that the indexical level implies, like in Peirce's examples (smoke-fire) a kind of homologous relationship between the sound of evocation and the sound production (the high 'F' of the 'Queen of the Night' is an 'index' of the physical/technical stress that the sound production requires) or, even, between the sound of evocation and the sound of reference (like in the cuckoo calls in Beethoven's 6th Symphony).

---

[32] Other theories draw also on the ideas of Peirce's semiotic: Clarke (1995:27), for instance, exploits, for purposes of performance expression, the same set of concepts (Index, Icon and Symbol) to serve as the basis for a more inclusive theory of expression. (The analytical framework concerned with the perceptual link between musical structure, expressive musical performance, and motion, in Chapter 3, also introduces some nuances concerning the interface between the different semiological levels).

The *iconic level* represents the sounds and patterns of a musical structure, or the dynamics of a musical line. Dowling & Harwood (1986:206) claim, drawing also on Langer's (1953) conception of music mimicking the form of emotional life. This level of music emotional evocation is founded in Dowling and Harwood's account through an important review of the empirical research existent on the interface between emotion and the music structural features (similar to the empirical evidence that is covered in Chapter 3 of this thesis under the item 3.3 'The expressive musical event: A semiological review of the empirical evidence of the expression of emotion in music/performance'). [33]

The 'symbolic level' presents a concept of emotional evocation very near to Meyer's (1956) account of music embodiment of emotional meaning. At this level, sounds might evoke an emotion in the listener, by reason of their place and function within a musical structure. As in Meyer's (1956) account, the clearer the listener understands the musical idiom(s) in question, the more available this level of representation becomes in and of itself. The tripartite distinction (Index, Icon and Symbol) well reflects the implicit distinctions made by empirical research in music and emotion field, because they fit well onto the identified categories of emotional response to music. Thus, and according to the 'semiological model', the empirical work taken in the field of music and emotion presents mainly three different lines of investigation:

- Studies of the way in which listeners associate musical works with extra-musical concepts (indexical associations).
- Studies of the various parameters that affect 'emotionality' of music structure (icons).
- Studies of musical structure and its ability to evoke emotional responses in the acculturated listener (symbols).

---

[33] Indeed, the iconic level of music emotional evocation is actually, a relevant conceptual field for the empirical research. Bierwisch's (1979) work, for instance, explores the possibility of the existence of a homology between the sounds of music and forms of the emotional life. As Bierwisch (1979:62) argues, "a sound pattern which is supposed to show excitement has to be excited." It is also the case of Knepler's (1977:126) concept of 'tuning element': "the origins of acoustic elements through which beings tune into one another reach back to pre-history... to pre-linguistic, pre-musical and pre-human times."

Nevertheless, and despite the fact that the tripartite categorisation fits closely into the empirical research, it seems that this categorisation does not apply so well when considered in the light of the phenomenon of emotion response itself.[34] Indeed, however exhaustive an analytical approach might be, it seems that it is always difficult to reduce reality even by the most detailed of the detailed theoretical distinctions. In fact, in many circumstances, the empirical evidence that supports these theoretical distinctions show that they are artificial at best. Often, the theoretical distinctions are just an indispensable artefact for purposes of the investigation.

In terms of the music performance practice, all these levels and sub-levels come across, most of the time, as mixed phenomena, ignoring thus all the theoretical distinctions. For example, it seems that some musical structures, usually experienced as a part of the 'indexical' and of the 'iconic' level of the expressive musical event have become through a process of sociocultural conventionalisation a part of the 'symbolic' level of experiencing and perceiving the emotional meaning of a music (see Kappas et al., 1991; Gabrielsson, 1991; Sloboda, 1991, 2000). As Lavy (2001) argues:

> "It is the very fact that certain structural units in Western tonal music have become *associated* with concepts, moods and attitudes that leads to their employment in anthem music in the first place. Such a suggestion, of course is right at the heart of Cooke's (1959) theory of musical language."

(Lavy, 2001:64)

---

[34] Lavy (2001 9) points out that: "Not only does this difficulty of placing Meyer's work and the empirical evidence that supports it show the distinction between icon and symbol to be artificial at best, but also music theory and plain common sense conspire to indicate that symbol and index can often been interchangeable. He argues that, "a perfect cadence, which is both highly recognisable and conventional musical unit, is certainly part of the symbolic structure – in itself encapsulating a complete expectancy generation and resolution cycle – but also has the potential to act indexically because as a tokenistic, recognisable, closed unit, it could well spark associations both cultural and idiosyncratic for a listener; it seems plausible, in fact, that in many circumstances, emotional response to musical structures such as perfect cadences could be understood much better as a phenomenon allied to indexical associations that to one allied to the tensions and resolutions of a musical dynamic." In short, it seems that the distinctions are, indeed, merely an artefact of theorising, and depending on the perspective music may assume one or the other, or even both semiological coding (for further detail see Lavy, 2001).

Indeed, Cooke's theory is the extreme case of a theory based in the idea that, a perceived associative relationship between musical features and emotions, mostly culturally contingent and arbitrarily found, evolved through a process of historically socio-cultural conventionalisation into a symbolic, quasi-denotative relationship between the musical structural units and the commonly associated emotional meaning. Other example might be found in some of the contingent and arbitrary associations culturally formed between film-music pairings (Sloboda, 2001:95). Indeed, the borders and the distinctions between icon and symbol, and symbol and index, are sometimes so small and diluted in music semiotics that these categories may become easily overlapped, and the distinction between symbol and icon level functioning difficult to detect and to distinct. Clarke (1995:28) argues, for instance, when referring to music performance expression, that:

> "There is no hard–and-fast distinction between iconic and symbolic functions, only different degrees of arbitrariness or convention. Performance expression appears to lie towards the iconic end of this continuum, but it is not entirely devoid of a symbolic component. The very minor role of symbolic signification suggests that performance expression is a rather unsystematic semiotic domain in which the more primitive indexical and iconic functions are paramount, and in which much of its contribution to musical meaning must be understood with reference to the human body."

(Clarke, 1995:28)

The semiotic theoretical account however is not the only framework to consider the relationship between music and emotion within a model of emotional response to music. Another relevant model bases its emotional response to music approach by placing the weight of this interface in the listener.

B) A LISTENER-CENTRIC MODEL

Lavy (2001:11) considers that, the point of departure for an empirical framework for research on emotion and music has to be situated "within a model of emotional response to music at the heart of which is placed the listener." Such a model, Lavy claims:

"must avoid arbitrary categorisations based on *a priori* concepts and musicological or philosophical assumptions; it must not see response to music as a phenomenon that is somehow special and isolated from other facets of human emotion. Instead it must seek to situate what we know about emotional response to music squarely within the frame of everyday experience."

(Lavy, 2001:11)

Thus, according to Lavy, emotional response to music should be centred in the *listener*. And, though the listener's responses to extrinsic sources of emotion are not exclusively musical, they are nevertheless intrinsic of the listening process.[35] So, instead of the question what are the features of music that evoke emotional responses in listeners? Lavy (2001:11) asks: "what are the features of the listeners that make them interpret music as an emotional stimulus?" So, what he proposes as a framework for empirical research is:

- "The foundation of a model of emotional response to music at the heart of which is placed the listener";
- "Such model must avoid arbitrary categorisations based on a priori concepts and musicological or philosophical assumptions;"
- "The proposed model should not see response to music as a phenomenon that is somehow special and isolated from other facets of human emotion; instead, it must seek to situate what we know about emotional response to music squarely within the frame of everyday human experience."
- Instead of attempting a categorisation of music itself, the model will be found upon four basic premises concerning listeners and their relationship to music and emotion, which are:

    1. Music heard as a sound: Just like any other stimulus in the auditory environment, music exists to be monitored and analysed.

---

[35] The ontological foundation of an 'intrinsic process of listening' might be seen as Lavy's contribution to an way out of the 'intrinsic vs extrinsic' musical meaning quarrel, a corollary of which is, in Lavy's (2001:10) words, 'the assumption that, emotional responses evoked by extrinsic sources and those evoked by intrinsic sources are quite different phenomena.' So, it seems that the real problem underlying the referred dichotomy seen in the empirical and theoretical work on music and emotion might be connected with the generally accepted assumption that intrinsic and extrinsic sources of emotion in music are ontologically distinct entities.

2. <u>Music heard as human utterance</u>: Humans have a remarkable ability to communicate and detect Emotion in the contours and timbres of vocal utterances; this ability is not suddenly lost during a musical listening experience.

3. <u>Music heard in context</u>: Music is always heard within the context of a complex web of knowledge, thoughts and environment, all of which can potentially contribute to an emotional experience.

4. <u>Music heard as narrative</u>: Listening to music involves the integration of sounds, utterances and context into dynamic, coherent experience. Such integration, far from being a phenomenon specific to music listening, is underpinned by generic narrative processes.

(Lavy, 2001: v-vi)

The importance of a listener's centric model for empirical investigation can be seen in the insight to face music as an essential part of everyday life, avoiding all metaphysical and dualistic conceptions of the musical phenomenon by shifting the focus of the research from the idealistic 'musical object' to the ecological level of everyday functional listening.

At the other end of the musical performance communication - the performer - might be considered as a pivotal element in the elaboration of a theoretical account based on the study of the perception of emotion. The performer's centred model - a counterpart of the listener's centred model - will be now considered as another important theoretical framework, and a relevant perspective to the study of the music-emotion's interface.

2.2.3 MUSIC AND THE COMMUNICATION OF EMOTION IN MUSIC/PERFORMANCE (SINGING)

A FUNCTIONALIST MODEL

A theoretical framework elaborated specifically to the study of the communication of emotion in music performance centred in the performer has been presented by Juslin (1995, 1997, 2001). He suggests that researchers should adopt a functionalist perspective on emotional communication in music performance, which involves the integration of ideas from research on emotion and non-verbal communication with Brunswik's (1956)

probabilistic functionalism. The functionalist perspective in music performance considers that the functions that communication of emotion has served in social interaction (Darwin 1872) must still be present and constraining the psychological mechanisms of emotion communication within the music performance interaction (Juslin, 1998). According to him, communication of emotion in music performance reflects two different factors:

a) The evidence for 'innate brain programs' (Jurgens, 1982)

b) The 'social learning' of emotional expression and interaction (Papousek, 1996).

These two factors, nature and nurture, Juslin (2001) hypothesises, are the origin of the acoustical code used by performers:

- Firstly, music performers communicate emotion to listeners by using the same acoustic code as in everyday vocal expression, a hypothesis supported by the similarities in cue utilisation between music performance and vocal expression, and by cross-cultural studies of vocal expression (Johnstone & Scherer, 2000) and music performance (Juslin, 2001). There is evidence of cross-cultural studies in India (Balkwill & Thompson 1999), Russia (Kotlyar & Morozov, 1976), Italy (Baroni & Finarelli, 1994), Japan (Ohgushi & Hattori, 1966b), Sweeden (Juslin, 1977) and Germany (Langeheinecke et al., 1999) suggesting that cues like tempo, sound level and timbre are used in a similar manner cross-culturally.

- Secondly, performer's emotional expression in music performance is governed according to social learning and specific memories, a life-long process, beginning with the socialisation of emotions. This hypothesis seems to be supported by the empirical evidence in emotional communication enlightened by the developmental research and infant studies of Stern (1985), Papousek (1996) and Trevarthen (1999). Thus, performers also learn how to link between acoustical cues and emotional expression, along with accumulated experience.

Juslin (1995, 1998, 2000) suggests that one way of capturing the crucial characteristics of the communication of emotion in music performance is to conceptualise it in terms of a modified version of Brunswiks's (1956) lens model.

Figure I (below) shows a simplified Brunswikian lens model for communication of emotion in music performance adapted from Juslin (2001:324). (For further details regarding information and methodology see Juslin, 1997; 2000; 2001; and 2002):
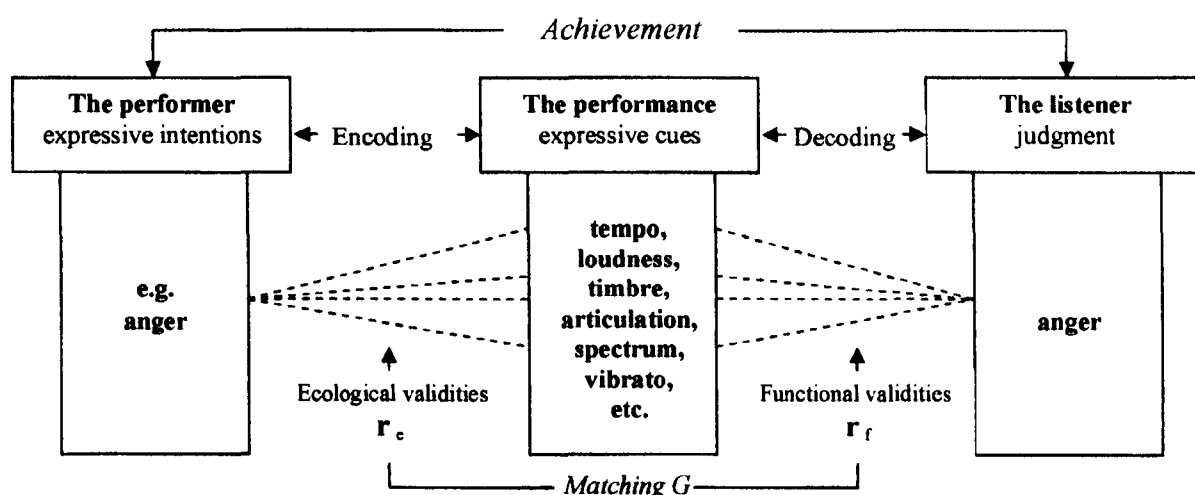
Figure I - Shows a Brunswikian lens model of emotional communication in music performance (adapted from Juslin 2001:324). Accuracy refers to the correlation between the performer's intention and the listener's judgement. Matching refers to the degree of similarity between the performer's and the listener's cue utilisation, respectively. For successful communication to occur, the performer's cue utilisation must be as similar as possible to the listener's cue utilisation. The performance expressive cues are: tempo, loudness, timbre, articulation, etc.

Originally conceived as a model of visual perception describing the relationship between an organism and distal cues, the lens model was later used in studies of human judgement relating the judge's judgement strategy to a description of the judgement task (Cooksey, 1996). Juslin's (2002:226) describes the modified lens model as follows:

"The modified lens model (...) illustrates how performers *encode* (i.e., express) emotions by means of s set of cues (e.g., variations of tempo, sound level, and timbre) that are *probabilistic* (i.e., uncertain) and partly *redundant*. The emotions are *decoded* (i.e., recognised) by listeners who use these same cues to infer the expression. The cues are probabilistic in that they are not perfectly reliable indicators of the intended expression. Performers and listeners have to *combine* the cues for reliable communication to occur. However this is not simply a matter of pattern matching, because the cues contribute in an additive fashion to listener's judgements: Each cue is neither necessary nor sufficient, but the larger the number of cues used, the more reliable the communication. The redundancy among cues partly reflects how sounds are produced on instruments (e.g., a harder string attack produces a tone that is both louder and sharper in timbre). [...] *Accuracy* refers to the correlation between the performer's intention and the listener's judgement. *Matching* refers to the degree of similarity between the

performer's and the listener's cue utilisation, respectively. For successful communication to occur, the performer's cue utilisation must be as similar as possible to the listener's cue utilisation."

So, the goal of such a framework is to provide performers, through empirical investigation, with information about their use of expressive cues in music performance emotional communication. The CFB (cognitive feedback) given to the performer by means of the Lens Model utilisation, allows them to check upon the accuracy of the expressive cues used and, thus, to improve the expressiveness of their performances (see Juslin 2000:151). Figure II (below) shows a summary of cue utilisation in performer's communication of emotion in music (see Juslin, 2001:315):

Positive valence

- TENDERNESS

slow mean tempo
slow tone attacks
low sound level
legato articulation
soft timbre
large timing variations

- HAPPINESS

fast mean tempo
fast tone attacks
higt sound level
staccato articulation
bright timbre
small timing variations

Low activity ←——————————————————————→ *High activity*

- SADNESS

slow mean tempo
slow tone attacks
low sound level
legato articulation
dull timbre
large timing variations

- FEAR

fast mean tempo
fast tone attacks
large sound level variability
staccato articulation
sharp microintonation
large timing variations

- ANGER

fast mean tempo
abrupt tone attacks
higt sound level
staccato articulation
sharp timbre
small temple variability

Negative valence

Figure II – Summary of cue utilization in performers' communication of emotion in music/performance (adapted from Juslin, 2001:315)

The performer's cognitive feedback, based on the audience's perception and recognition of the emotional meaning expressed in the music performance of music, leads inevitably to expose the relevance for the study of the relationship between music and emotion of an 'ecological model' based in the listener's awareness of the music performance.

2.2.4 MUSIC AND THE PERCEPTION OF E-MOTION IN MUSIC/PERFORMANCE (SINGING).

AN 'ECOLOGICAL' MODEL

In the first chapter, it has been stated that, as regards music perception, not only is music multifaceted in its nature, but also there are different levels of music awareness, which imply a certain plurality of music's expressed meanings. According to Clarke (1985), a listener is apt to hear three different types of events based in the information in the musical sound - the *performance event*, the *structural event* and the *expressive event*. The common thread passing through each of these levels of awareness is, according to Shove and Repp (1995:60), the perception of movement. This assumption is strongly supported by Bregman's (1990) approach to the perceptual organisation of sound, known as "auditory scene analysis". According to the "auditory scene analysis", music is built out of sound, and therefore the way listeners perceive music should obey the same innate principles of aggregation and segregation that allow them in real life, for instance, to identify if someone's footsteps indicate whether the 'maker of the sounds' is walking or running, becoming nearer or going away, in panic or calming down. Bregman explains:

> "Transformations in loudness, timbre, and other acoustic properties may allow the listener to conclude that the maker of a sound is drawing nearer, becoming weaker or more aggressive, or changing in other ways. However, in order to justify such conclusions, it must be known that the sounds that bear these acoustic relations to one another are derived from the same source. [...] This strategy of allowing discrete elements to be the bearers of form or 'transformations' only if they are in the same stream is the foundation of our experience of sequential form in music as much as in real life."
>
> (Bregman, 1990: 469)

Therefore, at an ecological level, it might be accepted that the source of this movement is the human performer, i.e., musical movement is human movement. As Shove and Repp (1995:61) point out, "the human performer is the ultimate source of the movement perceived". In other words, the performer's movements do belong to the music we hear and perceive, whether it is a *performance event*, a *structural event* or an *expressive event*. It seems, thus, that, contrarily to what has been stated for a long time now by different formalistic and absolutist accounts – "though performers must move to produce music,

such movements exist outside the music proper" (Epstein 1981:197) - there is no ontological gap between 'music proper' and music performance, or performed music.[36] So, as regards the perception of music/performance, the three different levels of the listeners' musical awareness considered by the 'ecological model' are:

- The *performance event*, which is connected directly to the articulatory movements of the performer, for example, bowing, tonguing, singing, and so on.

- The *structural event* belongs to a more abstract level of music perception. The listener is attending to a sequence of sounds – melody, cadential progressions, articulation of motives, phrases, themes, etc – by reason of their place and function within a musical structure, which is an indirectly consequence of the patterned articulatory movements of the performer. Certainly, as Meyer (1956) argues, movement within the musical structure in a particular direction creates an expectation for further movement in that direction (law of good continuation). But, a performer may certainly enhance that movement by emphasising notes that are of particular significance in the composition, or even minimise or deny the effect by deviation of the movement from stylistic expectations. As Shove & Repp (1995) argue:

> "Articulatory movements are sound-structuring movements. The motion one attributes to a succession of tones -- including its pacing, its character, even, its directionality - belongs first and foremost to the performer. That a listener reports hearing a sonic object in motion, rather than a performer, reflects the listener's perceptual attitude towards the musical event."
>
> (Shove & Repp, 1995:61)

It seems thus that it is the perceptual attitude of the listener that determines the level of awareness at which the musical event will be heard. The type of movement heard will be of course implied by that mode of listening and by the performer's intention.

---

[36] According to Shove & Repp (1995:58), Pierce is one of the few theorists who have acknowledged the crucial link between performance movement and musical movement: "I recognised that a performer necessarily renders musical movement with the physical movements of playing: blowing, tonguing, breathing, fingering, foot pedalling and the manifold gestures which in playing encircle these elementary movements" Pierce (1978:24; quoted in Shove & Repp:58).

- Last, the *expressive musical event*, corresponds, in Shove & Repp's (1995:60) definition, to "what Gabrielsson (1973a, 1973b) calls the 'emotion character' of music, i.e., patterns of movement whose general characteristics, and/or effects, are similar to bodily movement symptomatic of human emotions, moods or feelings."

2.3 SUMMARY

Having made a general review of several important theoretical accounts that might establish the basis for different empirical frameworks on the research of emotional meaning in music, it certainly will be useful to look now at some of the resulting empirical evidence. From the general review above presented, many useful points from the various theories seem to fit and contribute towards establishing an empirical approach to the expression of emotion in performed music. The ecological approach seems particularly interesting because it seems able to explain that the sense of motion and of e-motion in music performance is perceptually real (see Clarke, 2001), derived from the ways in which the acoustical and visual arrays specify their sources. Source specification is as relevant to the perception of the cultural environment as it is to the natural one, and as Gibson pointed out, symbols are not profoundly different from things. According to him, "no symbol exists except as it is realised in sound, projected in light, mechanical contact, or the like" (Gibson, 1966:26). So, it seems crucial to be able to identify the phenomena contributing to the production and perception of performance emotion. It is for this reason that the next chapter focuses on semiotics because it is a theory and method aimed at describing and labelling the different entities of the communication process and, also, it is capable of articulating the different levels of signification of the musical performance. Surely, a framework elaborated for an empirical approach on the study of the expression of emotion in music performance should consider primarily the level of musical awareness concerned with the *expressive musical event*, i.e., a framework whose primarily aim is the perception of those patterns of movement whose general characteristics are similar to bodily movement symptomatic of human emotions, moods or feelings. Paraphrasing Baily (1977: 330) such a framework should be able to perceive emotional meaning in music as a product of body movement transduced into sound. According to him, the performer's movements affect musical structures in an effective way, and the motor patterns implied in different kinds of music change according to the particularities of each music because they are the result of the interaction between the morphology of the instrument, the player's sensorimotor capacities, and the structural 'design' of the different music patterns. In other

words, an investigation concerned with the *expressive musical event* seems to require a level of musical awareness that considers the performer, and the performer's movements, directly or indirectly, as the 'heart' of the relationship between music (singing, in the present study) and emotion.[37] Thus, the performer, for the purposes of the current thesis, will be viewed as the threshold between the intended emotional meaning of the composer carried through the musical structure (structural features and expressive cues conveyed through the score) and the emotional meaning conveyed in the performance (structural features and the expressive cues conveyed through, directly or indirectly, the bodily movements of the performer), and perceived or/and evoked in the listener/viewer.

The performer is, of course, also, a listener. Indeed, according to Shaffer (1995:18), the performer is actually a "privileged listener able to provide detailed information on the sound pattern he or she expects to hear in the music." So, he/she is, in fact, someone who seeks to choose from the musical structure those features (structural and expressive) that, according to his/her own interpretation of the musical piece, will be properly underlining the intended emotional meaning of the composer. Therefore, the emotional meaning the performer intends to 're-create' will be communicated to the audience through the appropriated musical features produced by his/her structural and expressive bodily movements. In other words, in order to make this meaning emotionally clear and perceptible to the listener, and capable of inducing or provoking in the audience the intended emotional effect, the performer will have to re-create the appropriated musical features (structurally and expressively) through the movements of his/her body, in a way that the listener might understand, recognise and, even, feel the emotion conveyed. Thus, and with the purpose to become more familiarised with the study of the emergent issue of the 'expressive musical event' (the 'e-motion character' of music), and the role of the performer, a series of empirical studies and experiments, undertaken in music/ performance research will be reviewed from a semiotic perspective in Chapter 3.

---

[37] Nevertheless, an eye and an ear will be kept on the different nuances that the ensemble of the above referred theoretical frameworks might offer for more relevant empirical evidence on the subject of the expression of emotion in singing.

CHAPTER 3

3.   EMPIRICAL BACKGROUND

REVIEW OF SOME OF THE EMPIRICAL EVIDENCE ON THE EXPRESSION OF E-MOTION IN MUSIC PERFORMANCE (SINGING)

3.1   INTRODUCTION: SINGING AS A UNIVERSAL TRAIT OF HUMAN CULTURE

Chapter 2 notes that the role singing plays, varies greatly according to the different uses or ways in which it is employed and that a set of different functions may appear at every occasion. But, because the emotional content in singing and the communication aspect of its performance seem still to have nowadays a considerable relevance for singers and their audience, it seemed important to focus the research of this thesis on the factors and emotional features that intentionally (consciously or unconsciously) appear to express emotional meaning and, often, to evoke and arouse emotional response.

From an anthropological point of view singing has been considered as a universal form of human behaviour. The two major conclusions drawn from the work of Lomax (1968: vii, ix, xi), the 'Cantometrics Project', have considered:

1. Singing is a universal form of human behaviour:

"Singing is a universal human trait found in all known cultures as a specialised and easily identifiable kind of vocal behaviour. Singing is a universal trait of culture."

2. Song is an integral part of culture:

"Song style symbolises and reinforces certain aspects of social structure in all cultures. In general, a culture's song performance style seemed to represent generalised aspects of its social and communication systems."

The Cantometrics Project has developed a "coding book" to measure and categorise song styles from cultures all over the world in such a way that "the main families of song performance may be recognised, their geographical distribution mapped, and their relationship to cultural continuity, acculturation, and the expressive arts perceived" (Lomax 1968:35). But, if nowadays, drawing on 'Cantometrics Project' results, it is

generally accepted, that music and singing are universal traits of human behaviour, the role of music and its function are not, for instance, in Merriam's (1964) opinion, the same for all people. In his book, he lists ten different functions of music that help to explain what different roles music might play for different societies and within any society: the *function of emotional expression*, the *function of aesthetic enjoyment*, the *function of entertainment*, the *function of communication*, the *function of symbolic representation*, the *function of physical response*, the *function of enforcing conformity to social norms*, the *function of validation of social institutions and religious rituals*, the *function of the contribution to the continuity and stability of culture*, and *the function of contribution to the integration of society*. According to Hodges & Haack (1999:485), the role that music plays varies greatly, and many of these functions may overlap and occur simultaneously within the same musical experience. Some of music's functions are well recognised since Ancient Times, for instance, the *function of physical response* (motional) and the *function of emotional expression* (emotional) which clearly overlap and simultaneously occur with the *function of communication*. [38] It is thus the empirical evidence revealed in the main context of the study of these complementary functions of music/performance that will be now reviewed.

3.2   EMPIRICAL EVIDENCE ON MUSIC/PERFORMANCE (SINGING) AND E-MOTION:

A SHORT REVIEW.

Since Hanslick's (1986) day the idea that music can be related to some sort of movement and that its motional qualities can be heard as imitating the dynamics of behaviour and the dynamic qualities of feeling and emotion has become an important topic of theorists and researchers. The understanding and identification of musical motion had a first and pioneer empirical approach during the early decades of the twentieth century through the work of Sievers (1924), Becking (1923-4) and Truslit (1938). Truslit, for instance, showed that different kinds of movement image given to performers resulted in measurably different

---

[38] In Chapter-2, it has been already noted how strongly the ancient Greeks acknowledged the close relationship between 'music' and 'emotion' (see, footnote 28). However, it must be also underlined that the strong relationship between 'music' and 'motion' was not less acknowledged. Aristoxenus of Tarentum in *Elementa Rhytmica*, c. 320 B.C. (Barker, 1989:186), makes a clear reference to rhythmical movement and its analogy to physical movement: "I am thinking of things moved in the way that voice is moved in speaking and in singing, and the body in making a gesture and dancing..." According to Sachs (1943, quoted in Todd, 1995), "the connection between singing and arm motion is so narrow in ancient Egyptians that they expressed the meaning to 'sing' by the paraphrase 'to play with the hand'.

performances. So, he concluded that a very reduced number of movement types seems to underpin the kinematic basis of musical performance (for a detailed explanation of these different empirical approaches to musical motion see Shove & Repp, 1995:64-72). Nowadays, the influence of Gibson's (1979) theory of perception and information made clear the need to reassess the concept of musical motion, and the relevance of its role in the music performance and musical meaning has stimulated the empirical research to investigate the connection between expressive body movement and music. The many different approaches to this issue tend to focus, according to Shove and Repp (1995), on four major topics: (1) the source of motion; (2) the organisation of motion; (3) the character, or quality, of motion; and (4) the listener's perception and response to motion.

The study of the visual component of expression in live performance is a much more recent field of research opened up by Davidson (1991, 1993). Her studies, also based in Gibson's (1979) ecological approach of perception, demonstrated that the body movements made by performers while playing contribute to the expressivity of the performance as judged by the audience. Davidson's demonstration seems to be mainly concerned with the detection of bodily movements as indicators of expressive intentions significantly linked to important music structural features provided in a notated score. Other authors, like Shaffer (1992:265), for instance, claim that performers, conscious or unconsciously, seem to use physical gestures associated with emotional states as a basis for shaping musical expression. Schmalfeld (1985:18), clearly speaks of a need to find the 'character' or 'drama' of the work within the structure as a mean to restore in the performance its dynamic quality. Ultimately, it is possible to consider and experimentally to confirm that, bodily movements functioning as indicators of expressive intentions linked to important music structures can also be seen as physical gestures given musical expression to emotional states presented through the music. Indeed, Bower (1992) argues that emotions play a significant role in learning, conceptualising, and performing music. And Persson (2001:279) claims that there is "a need for music psychology – and perhaps even more so for music education – to begin to understand 'the emotional frame of mind' that generates recall, certain motor patterns, and communicative potential."

Concurring to the research of expressive behaviour, Cutting & Kozlowski (1977) explored the nature of physical expression (in their work 'expression' mean the walker's gender and identity) and concluded that different parts of the body convey, even within a very short period of time, similar expressive information at different levels. For example, any two-second excerpt from an isolated body part such as the ankle would reveal information about gender and identity of the subject in question (Clarke & Davidson, 1998). Runeson & Frykholm (1983) demonstrated that covert mental dispositions are specified in movement and can be detected by observers. This evidence seems to concur to Johnson's (1987) view that our reality is shaped by the patterns of our bodily movement, the contours of our spatial and temporal orientation, and the forms of our interactions with objects. Johnson (1987) argues that the structural found in abstract domains is only meaningful because we experience structure in bodily movements and interactions with the environment, and metaphorically map those patterns onto experience in other spheres, interpreting it in their light. It seems that this strong relationship between bodily movement and meaning can also be seen as the bodily basis of musical meaning. Stubley (1966) affirms that the body's role in musical performance is more than a physical execution; rather, musical performance is 'something done through and with the body'. According to her, the performance of music situates the body in a musical field where the music's movement and that of the performer are experienced and perceived as an 'on-going tuning process'. In this sense, and from a psychological point of view, Lidov (1987) and Jackendoff (1988) have proposed that the somatic nature of musical experience and the sense of motion attributed to music are both due to the mediating effect of a body representation. Evidence from the empirical field of music performance research seems to corroborate these different accounts and theories. The research on ritards undertaken by Sundberg & Verrillo (1980), by Kronman & Sundberg (1987), Todd (1992), and Friberg & Sundberg (1999) suggest that ritards in music mimic the behaviour of a deceleration of physical motion as in walking or running. Todd (1995) argues that the sense of motion is a characteristic perceptual response to music. More recently, Todd (1999) adopts a neurobiological perspective and argues for a vestibular component, complemented by a sensory-motor element, within the sense of musical motion. Cox (2001:196) claims an *embodied musical meaning* based on the idea that overt or/and covert mimetic participation is fundamental to musical experience. According to him, 'overt forms' of mimetic

participation include toe tapping, dancing to music and singing along with the music, while 'covert forms' include sub-vocalisation and other aspects of motor imagery; both make part of the way we understand all of the overt gestures of music performance. In fact, musical performance involves specific motor actions, and the understanding of these actions as Gallese & Goldman (1998) suggest involves mimetic-imagery. Evidence from clinical studies measuring reflex activity, EMG activity, autonomic activity, and associated brain activity measured in PET scans and fMRI, suggest that we understand the movements, speech, and musical sounds produced by others in part via unconscious imitation of those we observe. These results confirm the development research and infant studies of Stern (1985), Papousek (1996), and Trevarthen (1999). As Trevarthen (1999) argues:

> "Time of the mind is expressed in movements of the body. The rhythms which all movements show give evidence of what may be called the different levels of the Intrinsic Motive Pulse or IMP. (...) The most intimate details of how we think and feel can be conveyed by the forms of expression that colour what we show and tell. (...) Patterns of expression become unforgettable and moving events, especially if we synchronise our appreciation of how they change, and sympathise with the emotion involved, for no expression of purpose and experience is devoid of emotional value, even if this emoting is just a matter of curiosity and investigative interest."
>
> (Trevarthen, 1999:157-213)

According to infant studies, behaviour of children is based in a *musical* interplay between mother and child. Infants imitate the vocalisations, facial expressions, and gestures of others around them and this gradual integration of the sensory-motor mechanisms of the body into higher levels of concept formation and metaphorical discourse is what has been referred as the notion of *embodied meaning* (Lakoff & Johnson, 1980). As Johnson (1987) explains: "The inferential structure of our abstract reasoning is a high refinement upon orderings in our bodily experience, a refinement that ignores much of what goes into our reasoning" (Johnson, 1987:5). But, infants not only imitate adults' behaviour, adults imitate infants' behaviour, too. This mutual imitation promotes mutual understanding, and the overt form of children's imitative behaviour can be seen as active also in adults through a similar process but which becomes less perceptible and is known as 'imitative covert behaviour' (Walton, 1997). Baddeley and Logie (1992) suggest that once spoken words

involve covert imitation it seems reasonable to expect that comprehension of sung words ought to involve covert imitation as well. In fact, as previously referred, studies of motor imagery (Gallese & Goldman, 1998; Fadiga & Gallese, 1997; Fadiga et al., 1998) report experimental evidence which suggests that understanding the observed behaviour of others involves imagining performing the same or similar actions. Fónagy's (1962, 1967, 1976) investigation into emotional speech reveals interestingly similar set of results. He demonstrated that there is a close correlation between visible and normally invisible body movements. By carrying out research on how much information of the facial expression a listener can infer just from listening to the voice, he found evidence of concomitant movements, both glottal and articulatory, equivalent of facial gestures. Every attitude is expressed by its own articulation (and glottal) pattern which reflects the mental contents of the attitude. According to Fónagy (1976), this could be interpreted as a materialisation of the state of emotion, a re-interpretation of it into a movement. So, it seems that every emotional attitude can be expressed by different sorts of movements, different modalities of communication. Sundberg (1987) also interprets the relationship between body movements, human emotion, and the physiology of voice production, as a translation of specific emotional modes into different expressive modalities: movement pattern and acoustical pattern.

These suggestions imply that every emotional attitude can be expressed by different sorts of movements, different modalities of communication. Clynes (1969, 1980) seems to corroborate these findings by arguing that phonatory and articulatory gestures are manifestations of a common expressive dynamic form. Clynes's (1977) "sentics" theory claims that emotion and its expression form a unit, an integrated system, and that an emotion ("sentic state") may be expressed by different output modalities, such as "gestures, tone of voice, facial expression, a dance step, musical phrase, etc" (Clynes, 1977). Thus, in other words, it seems that expressive body movements can be translated into acoustic terms in voice production, and vice versa. That is exactly what Salgado (2000, 2001, 2002) has shown by exploring the correlation between acoustic and visual concomitants of emotional expression in the singer's performance as perceived by an audience. Amstrong *et al.* (1995) seems to share a similar opinion as he, by comparing overt gestures and speech, writes that:

"(...) the difference is not in the form of production (both are articulatory movements of the body), but in the form of the signal. Some articulatory movements result in primarily acoustic signals. Others, including semiotic 'gestures' as well as natural signed language, result in primarily optical signals."

(Amstrong et al., 1995:45)

This and other empirical evidence will be now reviewed in much more detail using an analytical framework concerned with the perceptual link between musical structure, expressive musical performance and motion (here included e-motion) directly or indirectly connected with the expressive and interactive performer's bodily movements that articulate the musical sounds.

3.3 THE 'EXPRESSIVE MUSICAL EVENT': A SEMIOLOGICAL RE-VIEW OF THE EMPIRICAL EVIDENCE OF THE EXPRESSION OF EMOTION IN MUSIC PERFORMANCE (SINGING)

Recent investigations into musical performance (Clynes, 1977; Clarke, 1985; Sundberg, 1987; Sloboda, 1991; Davidson, 1991, 1993; Schaffer, 1992, 1995; Gabrielsson, 1995; Gabrielsson & Juslin, 1996; Imberty, 1997; Juslin & Laukka, 2000; Juslin, 1997; Meyer, 2001; Scherer, 1995; Scherer et al., 2000) suggest that different levels of awareness and perception are involved in the inference of emotion when attending music performance as an *expressive musical event*. Having this in mind, a general outline of an analytical framework was established, by reason of a semiological review of the empirical evidence of emotional expression in music/performance (singing).

A preliminary word about the conceptual foundation of the analytical framework here established should, however, be added. Firstly, the foundational conceptual background of this analytical tool comes mainly from four different founts: i) Peirce's (1931-8) semiotic work; ii) Nattiez's (1990) semiology of music; iii) Shepherd & Wicke's (1997) music semiological model; and iv) what constitutes nowadays the so called 'ecological acoustics' (complementary, according to Shove & Repp (1995), to Gibson's work in 'ecological optics'). Secondly, the other supporting conceptual work comes from a plethora of different analytical approaches to the empirical field of musical emotional meaning. (These

supporting concepts and the respective authors will be referred as the explanation follows). Meanwhile, a basic reference to the foundational conceptual background will be presented:

1. From Peirce's semiotic work, three basic distinctions will be taken. In what follows, it should be understood that from the three kinds of signs in Peirce's semiotic, *index* implies a causal relation between a sign and its object (as between smoke and fire), *icon* a relation of similarity (as between map and the terrain it represents), and *symbol* a purely arbitrary and conventional relationship, only by virtue of a system of signs of which it is a part (as a word and its meaning).

2. From Nattiez's (1990) semiology of music, the distinction between extroversive and introversive semiosis will be taken. The important point about introversive and extroversive semiosis in Nattiez's distinction is that they are not distinct processes of referencing and they are always tightly intertwined in actual musical experience. (Of course, as discussed in Footnote 33, all these theoretical distinctions and concepts have to be considered as an indispensable artefact for purposes of the investigation). Introversive semiosis implies both intramusical and intermusical referencing, and extroversive semiosis implies all that is extramusical referencing. In fact, these two ways of referencing could be seen as working complementarily to the understanding of musical meaning. The *introversis semiosis* could be seen as an extra-musical referencing that is implied by the intrinsic structure of the music, and the *extroversis semiosis* could be seen as a musical referencing that turns intrinsic to the music, the structure of all that has been in this way extramusically-referred. Indeed, as Bowman (1998:248) argues "one of the most distinctive qualitative features of much musical experience is the state of instability and dynamic fluidity that exists between these two modes of referencing."

3. From Shepherd and Wicke's (1997) performative semiological model, the three distinct potentials displayed by musical sound "in acting as a ground and pathway for the generation of meaning and significance in human societies "(Shepherd & Wicke, 97:205) will be taken. First, the sounds of music start acting in a mimetic way. They can imitate and so, in certain way refer to other sounds (musical or

extra-musical sounds). It is the case, as already referred, of the 'cuckoo calls' in Beethoven's the 6th Symphony or, the typical case of the musical domestication of 'brute' affect vocalisations (see pg. 110). However, the way these sounds refer is never denotative in the sense language is, because they refer to the characteristic sounds of phenomena and not to phenomena themselves. Thus, this form of reference "may be thought as 'denotative' in character, but mimetically and thus homologously so" (Shepherd & Wicke, 1997:155). Another case of this 'almost-denotative' character of musical sounds might be seen in the relationship between timbres and timbral elicitors or 'voice timbres and gender identities' (Shepherd, 1991) or, even, as said, in musical sounds directly connected to the articulatory movements of their production (bowing, tonguing, singing). Second, the sounds of music might start acting structurally "in an iconic fashion in evoking more directly the logics and structures of the inner life" (Shepherd & Wicke, 1997:158). In other words, the abstract character of the internal properties of sounds and the inherent properties of sounds in music means that "they have to function in a structural manner if they are to function at all in being implicated in processes of meaning construction" (Shepherd & Wicke, 1997:158). And third, musical sound might start acting as a 'notational device' in a way, which completely elides its inherent sonic qualities. Thus, "the established relationship between the musical sounds and the phenomena evoked or referred is not necessary, i.e., it does not need to carry any kind of inherent implication between the characteristics of the sounds and the characteristics of what is referred" (Shepherd & Wicke, 1997:206).

4. Finally, from the so-called 'ecological acoustics' account, it will be taken the notions of ecological validity and of relational invariants (An exposition of this account can be found in Chapter 1, pg. 62 and Chapter 2, pg. 92). According to Shove & Repp (1995:62), ecologically valid empirical research is based "on the study of natural encounters with one's environment, but also in the idea that the environment is rich in information about its structure and dynamics." It is also based in the assumption that "under favourable conditions, this information, collected during a natural encounter with the surrounding environment, is sufficiently available to the perceptual systems, thus enabling the perceiver to tailor

his or her actions to the environment." Gibson (1966) suggests that in order to perceive a stable environment, the perceiver must be able to detect constant, 'invariant' elements in the optical and acoustical arrays. In the same way, Gibson (1979:247) claims that the perception of a relatively permanent, yet ever-transforming environment, like music expressive motion, is dependent on the perceiver's ability to extract "the invariants of structure from the flux of stimulation while still noticing the flux". This extract of invariants does not have to be understood as a conventional measuring device. According to Balzano (1987), perception is in and of itself a 'species of measurement'. So, when perceiving music performance, there may be a certain number of fluctuating elements within the acoustical and visual array that could be extracted in form of 'invariants', enabling thus the perceiver to comprehend the (e-motional) meaning of the presented performance. Shaw & Pittenger (1978) argue that events are 'invariant patterns of change', and if two events are distinct it is so because they show different patterns of change. Shaw and Pittenger postulate two types of invariants: 'transformational' and 'structural'.[39]

---

[39] Shove & Repp's (1995:63) explanation of Shaw & Pittenger's (1978) invariants' distinction will be fully reproduced because of the relevance it might have for the understanding how the ecological account approaches, for instance, the meaning of the performance of singing: "Transformational invariants are those relational aspects of the information that specify the identity of a particular pattern of change. Whenever Kiri Te Kanawa sings, she regulates the moves of her vocal tract in a manner distinguishable from most other forms of human activity, even, speech. Consequently her finely controlled vocal movements structure the surrounding air, distributing the initial energy of her actions through a sound field. The sound spectrum and temporal course of the resulting acoustic wave map onto the degree and rate at which she physically changes pitch and modulates the force of air through her vocal tract. Thus one hears *singing*. Structural invariants, on the other hand, are relational properties specific to the structure of the source object undergoing a particular style change. For the event 'Kiri Te Kanawa singing', these properties will correspond to the invariant structural features of her vocal tract and any other part of her body which may affect her production of sound. Her unique physical characteristics lawfully constrain the manner she sings and, consequently, the sound she produces, thereby allowing us to distinguish her from other singers. (...) Though conceptually distinct, structural and transformational invariants are functionally interdependent. One perceives 'Kiri Te Kanawa' only as one encounters her singing, talking, laughing, walking and so on. Her identity is intertwined with what she does as a physical, dynamic being. And yet what she does – how she sings – is lawfully determined by the structure of her body." Hypothetically, the specific values found for the two sort of relational invariants of the given 'Kiri Te Kanawa' example could be ascribed to traditional psychoacoustic variables. In that case, it seems possible to establish a relationship between the difference of the two concepts of invariants presented (transformational and structural) and what has been called by Husson (1962), the distinction between 'timbre vocalique' (vowel timbre) and 'timbre extra-vocalique' (extra-vowel timbre) of a singer. 'Vowel timbre' is what distinguishes a vowel A from a vowel E within the same voice producer. So, using the given 'Kiri Te Kanawa' example, 'vowel timbre' is what will let us distinguish between Kanawa singing the word 'BAD' and singing the word 'BED'. Thus, 'vowel-timbre' is exclusively constituted by the harmonics near F_1 and F_2 of the selected vowel. 'Extra-vowel timbre', is what will let us distinguish between Kanawa and Callas singing the exact same vowels of the word 'Diva'. Thus, 'extra-vowel timbre' is not characteristic of the vocally produced vowel, but of the voice of the producer, of the invariant structural features of the singer's vocal tract (for further details concerning the difference between 'timbre vocalique' and 'timbre extra-vocalique' see Husson, 1962:67-70).

It has been said in the previous chapter that musical structures are affected in an effective way by the performer's movements and actions or gestures (movement's intention) and, also, that the motor patterns implicated in and by different kinds of musical structures change according to the particularities of each music. Firstly, by way of the interaction between musical structures, states of awareness, the player's sensorimotor capacities, and the morphology of the instrument. Secondly, because not all music can receive all possible 'intended' meanings and not all 'intended' meanings can be successfully invested in all musics. As Shepherd & Wicke (1997) argue:

> "Music is meaningful to individuals inasmuch as the inherent characteristics of its textures and structures are suitable to the investment in them of meanings – themselves coded and experienced through the rhythms, textures and structures of the body - that such characteristics may call forth."
>
> (Shepherd & Wicke, 1997:178)

Having this in mind and considering the amount of empirical evidence on the subject of emotional expression in music performance, it is possible to suggest different levels of coding (and de-coding) of the so-called 'emotional character' of music:

A. First, the *Indexical-qualitative coding*, through which the emotional character of music (the *expressive event*) might be partly directly connected to the perception of the articulatory and interactive movements of the performer, and partly directly connected to the acoustic structure of small segments of sounds (timbres, for instance). These articulatory and interactive movements are a consequence of the performer's capacity to express and communicate emotion within the bodily movements and gestures that articulate the music and interact with other elements of the performance (instruments, co-performers, audience) critical in the generation of emotional meaning. The generation of the acoustic structure of small segments of sounds is a direct 'transduction' of the articulatory movements of the performer(s) and is usually connected with the elicitation of bodily and behavioural manifestations of emotion. A good example can be found in

Scherer (1995:236), who claims, following Kainz (1962) and Wundt (1900), that "the use of 'Ah' and 'ahimé' in Baroque operas illustrates what has been called the domestication of 'brute' affect vocalisations." An extensive discussion of such affect vocalisations, or 'facial/vocal affect-bursts', can be found in Scherer (1994). He has focused on the functions of these affect vocalisations, and concludes that these serve practically all of the semantic, syntactic, pragmatic, and dialogical functions of non-verbal behaviour (Scherer, 1980). According to Shepherd & Wicke (1997:155), it seems possible to consider the sounds at the level of timbre functioning according to an indexical coding. They consider this possibility in relation either to actual sounds or to non-sonic phenomena. So, sounds may refer 'timbrically' to other sounds in the external world, as in the use of drums as a copy of the military drumming and to signal the concept of 'militariness', as in Puccini's Tosca (shortly before Cavaradossi's shooting), or even as a symbolic evocation, as in the case of the 'cuckoo calls' in Beethoven's 6th Symphony. Secondly, they can also evoke 'timbrically' various internal states, for instance, the already referred relationship between various voice timbres and gender identities (Shepherd, 1991) or, as in Tagg's (1991) example, the use of a distant and hollow quena flute sound to evoke the feeling of a 'longing' sigh. Nevertheless, Shepherd & Wicke (1997), consider this form of reference, as a special case of the homologous functioning of sounds in music:

> "Such sounds are not, however, capable of denoting actual phenomena in the external world: they refer only to the characteristic sounds of phenomena and not the phenomena themselves. This form of reference may be thought of as 'denotative' in character, but mimetically and thus homologously so."
>
> (Shepherd & Wicke, 1997:155)

Of course, the ecological account allows to open up the field of functioning of the sounds and to claim that the indexical coding may also include the reference to other levels of the investigated phenomena, such as the movements that articulate the production and the expressiveness of those sounds. This seems to be corroborated by Fónagy's (1962, 1967) investigation of emotional speech, and by Clynes's (1977) "sentics" theory. They claim that emotion and its expression form a unit where emotion can be expressed by different output modalities, such as gestures, tone of voice, facial expression, a dance step, musical phrase, etc. Thus, phonatory and articulatory gestures are manifestations of a common

expressive dynamic form which underlies both the perception and production of expression in different modalities (Clynes, 1969), and music performance as whole can be regarded as the ensemble of expressive gestures, sonic or articulatory, which are different manifestations of the same interpretative musical intention.

The indexical-qualitative coding works basically at a micro-structural level. Scherer & Zentner (2001:362) name these micro-structural acoustic characteristics of the building blocks of musical structure, *segmental features*. According to Scherer & Zentner (2001), these segmental features are expected to be relatively stable and universal on emotional inference. Among the ensemble of movements used to articulate structure and meaning at the indexical coding level it seems possible to find all those movements characteristics of expressive modalities and involved in emotional communication, particularly body posture, facial expression, and vocalisation. According to Clarke & Davidson (1998), the reason why certain segmental excerpts of the performance are more richly informative that others might be seen as a consequence of the use of certain mimetic gestures, and that these movements heighten the expressive impact of a specific moment in the music. In fact, Ekman et al. (1983) provided evidence that it might be possible to induce physiological and experiential emotion reactions by asking someone to produce the facial patterning that is characteristic of certain emotions (without the person being aware of this purpose). This idea is also consistent with proprioceptive feedback theories, such as facial feedback, that claim that subjective feeling can be produced or intensified by increased motor expression (see McIntosh, 1996). Lipps (1909) argued that understanding the emotions of others occurs through empathy - 'Einfühlung' - which is based on the human capacity of mimicking the expressive patterns seen in the others. This could be also explained as a sort of 'contagion' of the expressive moment. Eisenberg & Strayer (1987) call it 'motor mimicry', and it is assumed to be a process by which a person observing the strong motor expressions of another, even without any knowledge of the event that caused that reaction, will feel to produce in his/herself the same muscular innervation. Scherer (1992) pointed out the strong possibility of the existence of a link between vocalisation and induction of emotion through mimicry of perceived physical expression. Scherer (1995:235) claims that, "as far as facial expression is concerned, behavioural and social scientists have much progress in gathering evidence on the phylogenetic continuity, the

universality across cultures, and the rich information content of emotion in facial expressions". The work of Tomkins (1984) and of his followers Ekman (1984, 1992), and Izard (1971, 1994) reasserted the Darwinian notion of a limited number of 'basic' or 'fundamental' emotions serving purposes of adaptation and social interaction. So, it seems arguable that specific moments of the musical structure may elicit some of those basic emotions (and even non-basic emotions), and in that case the induction of the (basic) emotion would require the production of the specific patterns of physiological reactions postulated for them. Therefore, it might also be the case that some of the performer's movements that articulate specific structural segmental features re-produce by motor mimicry, or some other mimetic function, the physiological patterns, such as facial gestures, differential action tendencies, etc, which are concomitants of the expression and elicitation of (basic) emotions. The importance of vocalisation and facial movements as different modalities of emotional expression (as different manifestations of a common emotional expressive dynamic form which underlies both the perception and production of expression) makes the indexical-qualitative coding, especially interesting for the study of the communication of emotion in the performance of singing.

B. Second, the *Iconic-dynamic-introversive coding*, through which the expressive event might be a directly implication of specific and conventionalised structural and transformational features of the performed music 'symptomatic' of human emotional states. At the level of *Iconic-dynamic-introversive coding*, the *expressive event* can be perceived as being directly connected with the structure and transformation of the imagined 'musical object', and a direct or indirect consequence of the performer's structure and transformation. In other words, this embodied emotional meaning is an achievement of the dynamic sensations of flux, tension, and expectations fulfilled or violated, that are latent in the musical structure and that can be emphasised or frustrated, directly or indirectly, through the articulatory movements of the performer.

i) Music structure and 'introversive coding'

According to Meyer (2001), we classify emotions as we understand and respond to different styles of music, or as we classify the world in general. It depends on learning and internalising the norms that govern relationships and, consequently, comprehend the

implications of patterns, and experience deviations. Meyer argues that, "to choose effectively, it is necessary to imagine the probability of coming events (whether the behaviour of people or of musical patterns) and such imagining invariably entails classification" (Meyer, 2001: 343). In other words, an act of recognising and experiencing the relationships involved in an event, and a conscious classification and naming of the kinds of relationships, techniques and strategies presented within the dynamic and structure of the event. Nevertheless, empirical evidence has shown that the decoding of emotion from music performance is largely independent of musical training (Juslin 1997) and that even children, 3-4 years old, are consistently able to decode such basic emotions as happiness, sadness, anger, and fear (see Dolgin & Adelson, 1990; Kratus 1993). But, Meyer (2001:343) argues equally, that "awareness and response to the class of some phenomenon need *not*, however, involve conscious conceptualisation or naming, but rather recognition and experience of the relationships presented - that is knowing how the kind of patterning *works*." Therefore, it seems that, for Meyer too, the so-called music expressive event does not always have to involve conscious classification. Meyer's (2001:343) explanation is that, "enjoying music is a matter of experiencing tension of syntactic and native processes as well as the fluctuations of statistical parameters." Thus, it seems that what Meyer suggests is (also) a kind of human empathy with the action of the music that brings about analogous feelings in listeners. He considers two kinds of emotional entities - the emotional states and the emotional processes - both capable of generating empathy:

> "The distinctions between emotional states and emotional processes suggest that different kinds of relationships evoke and shape different facets of musical experience. The statistical parameters beget the physical-somatic conditions, which through empathy, characterise emotional states in primal present. Syntactic and native processes not only beget physical empathy with music, but generate mental arousal that, directing attention to what is still to come, gives rise to implication."
>
> (Meyer, 2001:346)

Meyer's distinction between 'statistical parameters' and 'native and syntactic processes' seems to fit well with the distinction proposed here between 'extroversive' and 'introversive' iconic-dynamic coding, as long as an interchangeable relationship between them might be possible to be postulated, as the one between the introversive and the

extroversive function of Nattiez's (1990) 'semiosis' distinction. In fact, Meyer seems to suggest this very same idea when he argues that:

> "The very existence of sound entails the use of statistical parameters. Consequently, native and syntactic processes are invariably coordinate with, and qualified by, the action of statistical parameters. Responses to these parameters are also native, rather than learned. But because they are essentially matters of amount, statistical means do not give rise to functional differentiation. Rather what they involve is the continuation of prevalent action. There is however, an important exception: when amounts – whether of information or redundancy, of speed or intensity of stimulation – become excessive in relation to human cognitive capabilities, the native implication is change rather than continuation. "
>
> (Meyer 2001:345)

To conclude and following closely Meyer's (2001:344-45) own words: Native and syntactic processes are both subject to cognitive constraints of the mind and to the physical ones of the body. Both give rise to implication and almost always complement one another in the shaping of musical experience. Nevertheless, while 'syntactical processes' involve learning the norms of a particular musical style and are therefore historically and culturally circumscribed, 'native processes' are less specific, and essentially universal. Statistical parameters, however, allow the identification and naming of emotional states as a result of the association of different dispositions of music's acoustic components with feelings characteristically evoked in natural or cultural circumstances. Other authors present a similar conceptualisation in what concerns the establishment of analytical tools in order to deal with this specific amount of empirical evidence on the subject of music and emotional meaning. Sloboda and Juslin (2001), for instance, present, in what concerns the different sources of emotion in music, a similar distinction between intrinsic emotion and extrinsic emotion. Sloboda and Juslin suggest that:

> "there may be a partial decoupling between the mechanisms that determine intensity of affect and those that determine emotional content, the former being predominantly determined by structural characteristics of the music (*intrinsic emotion*), the latter being determined more strongly by contextual factors, including the memories, associations, and priorities of the person hearing the music (*extrinsic emotion*). (…) The term intrinsic seems appropriate because it is only to reference to other musical events and structures that particular events become more or less expected. These expectations may reflect learning, or

may operate at the level of primitive perceptual processes, such as the 'gestalt laws' of perception. (...) Theory and evidence suggest two sources of content – iconic and associational. Because they both refer outside the music, we call these sources extrinsic."

(Sloboda & Juslin, 2001:92)

Based on empirical evidence by Goldstein (1980), Sloboda (1991, 1992) and Panksepp (1995), Sloboda & Juslin (2001) claim the existence of a lawful relationship between the musical discourse and the emotional experience:

"There is a lawful relationship between the intensity of emotional qualities experienced in music and the specific structural characteristics of the music at a particular point in time. The intensity of emotional response to a piece of music often rises and falls, as the music unfolds. Musical discourse, both formal and informal, talks of climaxes and points of repose, tension and relaxation. In other words, there are peaks, where intense emotions (or other affective sensations) are prone to be experienced, and troughs, where the intensity is weak."

(Sloboda & Juslin, 2001:91)

In fact, it seems, according to Gabrielsson (1995) that, the 'emotional character' in music often changes within a piece more or less gradually or, even, suddenly (Gabrielsson, 1995:38). An accurate experimental approach to the proposed relationship between the intensity of emotional perception and response to music and the musical discourse itself is the use of the so-called continuous (or 'continual', see Schubert 2001:394) recording of emotional expression. According to Gabrielsson & Lindström (2001:226), the type of the emotional expression also varies during the course of a musical piece. So, the use of a continuous recording technique of the perceived expression may also help for a detailed study of how factors in the musical structure affect the perceived expression. The first known attempt to a continuous evaluation of perceived emotion was Hevner's (1936) adjective checklist early exploration (self-report of a continuous emotional response approach to a musical piece). Later, Nielsen's (1983, 1987) pioneering research showed that "tension peaks were mainly conditioned by high intensity (fortissimo), but increased tension could also be related to ascending melody, increased note density, dissonance,

harmonic complexity, rhythmic complexity and formal properties such as repetition of various units, condensation of musical material, sequential development, and pauses."[40]

Gabrielsson (1995:38) commented on Nielsen's (1983) study arguing that he came close to studying the types of more global and time-varying properties suggested by Langer (in Nielsen's case tension and release). According to Gabrielsson & Lindström (2001:243), Langer's examples represent "complex combinations of perceptual, cognitive, and emotional aspects that elude conventional psychological terminology." In fact, Langer (1953:27; 1957:228) suggested that, by reason of the formal iconicity between music and human feeling, we may perceive 'expressions' such as growth and attenuation, flowing and stowing, conflict and resolution, excitement and calm, or 'patterns' of motion and rest, of tension and release, sudden change. Nowadays, a commonly shared idea to several music researchers (Imberty, 1997; Gabrielsson & Lindström, 2001; Sloboda & Juslin, 2001; Bunt & Pavlicevic, 2001) accepts these 'complex combinations of perceptual, cognitive, and emotional aspects' of music's expressive qualities related to structural changes across time, as phenomena that remind Stern's (1985) 'vitality affects'. 'Vitality affects'is a well-documented concept drawn from infant studies (Stern, 1985; Papousek, 1996; Trevarthen, 1999) and refers to a specific music-behavioural communication relationship or 'attunement' between mother and infant. Stern coined this concept to refer to specific qualities of the affective communication between parents and infants. He argues that the infant's entire world is one of "shapes, intensities and temporal patterns" and that these 'qualities' are a-modal and dynamic forms of feeling that may occur together with, or in the absence of, proper emotions (Stern, 1985:57).

An example of how the concept of 'vitality affects' can be applied to a phenomenological description of music experience can be found in Imberty (1997). He argues that musical expression might be seen, as a dynamic interactive process resulting from the 'proto-narrative' character of the musical time organised according to 'an architecture of vitality affects'. In one sophisticated attempt to investigate how emotional expression is

---

[40] Other studies of perceived tension followed, with other devices, made by Madsen & Fredrickson (1997), and Krumhansl (1996), who found that tension peaks followed by a rapid decrease occurred at the end of large-scale sections. Continuous recording of perceived emotional expression using computerised techniques was also researched by Krumhansl (1997), Madsen (1997) and Schubert (1999) (for further and extensive review, see Gabrielsson & Lindström, 2001, and Schubert, 2001).

influenced by properties of musical form, Imberty (1979) came to the following conclusions: a) low complexity combined with average dynamism means formal integration and expression of positive emotions; b) high formal complexity combined with low dynamism means formal disintegration and expression of melancholy and depression; c) high formal complexity combined with high dynamism means formal disintegration and expression of anxiety and aggressiveness (for further details, see Imberty, 1979).

Thus, it might be said that, enough empirical evidence from performance research field seems to concur on the idea that music is able to articulate significance and emotional affect in a pure and dynamically structured manner. Sloboda (1998:27), for instance, reports how particular structural musical features impacting on us are capable of affecting directly the materiality of the human body, provoking a reaction whose significance has been 'translated' by the individuals subjected to the experiment into tears, shivers, gooseflesh and racing heart. In one of his experimental studies, Sloboda (1991, 1992) used a retrospective questionnaire method in an attempt to associate specific structural elements in music with the reported occurrence of the 'thrill' emotional response, as indicated by physiological effects such as shivers down the spine, tears and racing heart. Musical analysis of those segments which participants could clearly pinpoint as able to evoke in them a significant emotional response revealed an interesting correlation. According to Sloboda's results:

> "Shivers were associated with enharmonic changes and other harmonic, textural, or dynamic discontinuities. Tears were associated with melodic or harmonic sequences, appoggiaturas and suspensions, and downward harmonic movement through the circle of fifths. And racing heart was associated with syncopation, and other forms of accentual anticipation. All of these responses were considered "pre-aesthetic"
>
> (Sloboda, 1998:27).

In another empirical study, Panksepp (1995) examined the "shivers down the spine", a form of physiological reaction to music in a rock/pop ballad style, and named it the 'chill' response. He found certain types of acoustic dynamics, such as *piercing crescendos* were necessary triggers for the 'chill' phenomenon. These two experiments show that, by regarding music performance as an expressive musical event, the listener hears these specific structural elements (patterns implications, realisations, etc.) of the performed

music as an emotionally charged gesture. So, musical performance could be one in which the performer "recognises the dynamic implications of the music structure and enhances the acoustic signal in such a way as to boost or highlight those implications" (Sloboda, 1998:25). Sloboda proposed the term 'analogical mapping' for naming the dynamic process of meaning construction by which the performer recognises the implications of the musical structure and performs them through articulatory movements and gestures that enhance the acoustic signal in such a way as increase their value, strength and significance. This seems also to be the dynamic process by which listeners, recognising the articulatory movements and gestures of the performer and the implications and significance of the musical structure, become aware of the emotional meaning of the *expressive musical event* (see Davidson, 1991, 1993).

ii) Music performance and 'introversive coding'

Enough empirical evidence exists, also, to support the close relationship between musical structure and expressive musical performance. Different perspectives, models, codes and approaches have been made during the last thirty years, which show a clear correlation between musical structure and the application of expressive devices. This research has been particularly concentrated on specific expressive musical factors, such as: i) 'Timing' properties (e.g. Clarke, 1988 - a study where he identified the use of timing, dynamics and articulation as means to convey the proper structure, and emphasised that they may substitute for one another or be used in combination depending on the underlying musical structure; Sundberg, 1988; Todd, 1985, 1988; Gabrielsson, 1982, 1988; Shaffer 1984; Rasch, 1988); ii) 'Dynamic' variations and musical structure (Clarke, 1988; Todd, 1992; Kendall & Carterette, 1990); iii) as 'Metre' (Sloboda, 1983); iv) 'Tempo' (Wagner, 1974; Repp, 1994; Desain & Honing, 1994; v) 'Articulation' (Clarke, 1988); vi)'Ritards' (Sundberg & Verrillo, 1980; Kronman & Sundberg, 1987; Todd, 1992); vii) 'Expressive markings' (Shaffer, 1995). Further and detailed information can be found in Gabrielsson's (1999) general review on the subject.

Although the study of timing, articulation and dynamic has dominated psychological research into the generation of expression, some authors have also developed studies into other parameters of musical performance expression. Sundberg (1995, 1999), for instance has extensively reviewed studies into aspects of vocal timbre, vibrato and pitch, formant

frequencies and intelligibility of vowels. Iowa researchers (Tiffin & Seashore, 1932) studied extensively singers' and players' use of vibrato. All these acoustic phenomena have been studied as part of the range of artistic deviation that singers employ to generate expression in music performance.

The use of physical gestures, consciously or unconsciously, to enhance information about the music structure and to demonstrate the performer's expressive intentions has been the core of Davidson's (1993, 1995) work. Davidson showed that different expressive performances of the same piece (without expression, normal expression and exaggerated expression) could be clearly perceived and differentiated even by listeners who have just watched them on video (only visual information). Other findings or results of Davidson's work showed that different regions of the body convey similar expressive information of the performance intentions. Some are more global indicators of expression (swaying motion as a more global idiomatic and kinematic character of music) while others provide much more local information (for instance specific information limited to just two seconds of the performance). The latter revealed that some moments of the performance were more obvious indicators of expressive intentions than other moments, and therefore that a significant link between the detectable intentions of the performer's expressive gestures and musical structures could be found. This suggests, also, that the most obvious indicators of expressive intention coincided with important structural moments (a cadence point, for instance). Further conclusions may be taken from these results: key musical structures seem to be the individual points around which expression of the intention is most accentuated. According to Clarke & Davidson (1998):

> "Given the strong evidence that structure and expression are closely related, significant structural moments are likely to provide the focal points around which specific examples of expression will be organised, accounting for the very local nature of some expressive moments."

> (Clarke & Davidson, 1998: 79)

The work of Davidson brings strong evidence to the idea that there is a close connection between the performer's expressive movements and the musical structure. Ultimately, it is plausible to consider that these bodily movements might be connected to those specific

features of the musical structure, which have empirically proved as being 'symptomatic' of human emotional states. For instance, they may emphasise and be intentionally expressive of the deep emotional meaning of a specific musical structure implication that they physically articulate. The way the iconic dynamic introversive coding functions in music performance seems comparable to what Hatten's (2001:8) calls 'embodying sound':

> "The performer explores, gesturally, the background coherence and foreground configuration of a work and to sense kinesthetically the pulls and releases implied by organic phrases in their unique combinations of tonal rhythmic forces. Ultimately, the performer learns to translate gestural character into sound, through the medium of the body's corresponding (intermodal) gestural realisations. Learning how to perform is thus inseparable from learning how the piece is structured, how it has expressive meaning, how one can physically manifest that meaning in one's body, and how one can then transfer that bodily gestural meaning to the instrument."
>
> (Hatten, 1999:L3:1)

Hatten's account on the role of movement in performance draws on Pierce (1994) pedagogical work. According to Hatten (1999:L3:1) what Pierce offers to her students is a nuanced pedagogy that accounts for the expressivity of musical processes, both those that issue straightforwardly from the style, and those that strategically deflect the interpreter from such a straightforward reading. So, this distinction accounts for two levels of intentional gestural meaning in music performance: one that accounts for the stylistic level of musical meaning, the other accounting for the strategic level of musical performance. The goal in Pierce's pedagogy is to achieve in sound the expressivity and implied meaning the student has previously explored and experienced through the embodied analytical exercises. In this sense preliminary exercises like 'arm swings', 'contouring', 'stepping' give the student the possibility to internalise, to embody essential aspects of the stylistic musical structure and prepare the student to the final stages of preparation which involve internalising of intentional curves and climaxes that shape the musical meaning, and embodying the peculiar character or tone that carries the music's implied emotional content. The two essential expressive attributes comprised by the final stages of preparation may be practiced separately as "spanning" and "tone of voice" (Pierce, 1994: 87-101). "Spanning" may be embodied in two ways by 'stretching' or by 'arcing', and it is considered by Pierce as the most purely Schenkerian of the movement processes. "Tone of

voice" is Pierce's term for "the wash of affect in a piece of music" which "suffuses the sound with the hue and the player with the expressiveness visible through the entire body" (Pierce, 1994:96). As Hatten points out they are two intermodal sources to understand the musical gesture: the physical movements of the body and the intonation curves of human utterance. According to him:

> "Tone of voice draws on the latter (intonation curves of human utterance) for its inspiration, but an intonation cannot be achieved without an intention, a semantic and a pragmatic context (...) The student should thereby transfer not merely kinetic energy but a particular quality of energy to his performance."
>
> (Hatten, 1999:I.3:2)

It seems thus arguable to establish a correlation between 'spanning' and *introversive dynamic coding* and another between 'tone' of voice and *extroversive dynamic coding*. It is to this proposition that we now turn.

C. Third, the *Iconic-dynamic-extroversive coding*, through which the expressive event can be directly connected to a 'virtual/potential'[41] source of emotion elicitation emergent from the virtual "trompe l'oreille" of an intended grouping of sounds dependent on the sound-structuring sequence generated by the performer's articulatory movements. This level of awareness seems to be inhered from a mechanism of perception and transmission (coding and decoding of information) of acoustic and visual cues used in emotion inference from vocal and facial communication. According to Bregman (1990):

> "A composer may want the listener to group the sounds from different instruments and hear this grouping as a sound with its own properties. This grouping is not a real source, however. Stephen Adams has referred to it as a 'virtual' source, perhaps drawing the analogy from the virtual image of optics. [...] Experiences of real sources and of virtual

---

[41] The term 'potential' seems to fit more accurately to the idea of a source of emotion that though intended to be perceived in a virtual way, may become real if the performer and the performed musical sound-structuring sequence creates a perceptual effect that will be experienced as a 'lived' emotional state. This is, recognised by many performers as a goal of the authentic and artistic musical performance. As Persson (1995) argues, musicians are 'sensors of emotion and sensitivity', and it will be helpful for future research "to term the dynamic nature of this emotional basis from which musicians appear to draw motivation, construe artistic understanding, and generate performances, a *musical reality*." Persson (2001:282) reports how musicians "consciously manipulate recall of certain memories", i.e., 'emotional memories', *mood induction*, or/and 'memories of emotions', *imagery*, in order to 'generate' a particular emotion.

> sources are both examples of auditory streams. They are different not in terms of their
> psychological properties, but in the reality of the things that they refer to in the world. Real
> sources tell a true story; virtual sources a fictional."

<div align="right">(Bregman, 1990:460)</div>

As Clarke (2001: 221), also, explains:

> "Music may create perceptual effects with the disposition of discrete pitches and
> instrumental timbres in time that reproduce, or approximate to, those that we experience
> with the continuous acoustical transformations that are characteristic of real world events."

<div align="right">(Clarke, 2001:221)</div>

The initial investigation in expressive musical performance was convincingly able, to establish a close correlation between music structure and expressive musical performance and, also, a set of performance generative rules and models that were able to predict in a fairly accurate way the expressive variations and deviations used. In the referred approach, however, the primary, if not the 'only' source of performance expression that was taken in to consideration was one directly connected with a mental representation of musical structure. This conception is clearly inherited from a more general structuralist conception that considers structure as the only possible 'referent' in a chain of referents and interpretants (see Chapter 1, 'The structuralist Contribution'). However, the idea that structure alone generates musical expression in performance has revealed itself through the years insufficient to establish an accurate explanation to all the more practical and corporeal nuances of an expressive musical performance. Thus, the theoretical and empirical research had to start looking further, upon other factors that might explain the generation of expressive performance in a more complete and realistic perspective. Clarke and Davidson (1998) seem to corroborate this view when they argue that:

> "a model which portrays performance simply as the flow of information from input through
> a set of abstract expressive rules to an output effector system is implausibly abstract and
> cerebral, and the reality is far more practical and corporeal. The body is not just a sort of
> sensory input and a mechanism for effecting output: it is far more intimately bound up with
> our whole response to music – perceptual and motor – and needs to be recognised as having
> a far more central role than simple a generative model would suggest. A wide range of
> other factors, including the possibilities of the instrument, the acoustics of the performing

environment, the nature of the audience, the mood and intentions of the performer, and even the performance ideology espoused (e.g. 'historically aware' performance practice) will contribute to the result, sometimes to the detriment of structure (as in case of an indulgent or wayward performer), but preferably in accordance with an interpretation of the structure. Movement, and the human body, are particularly significant in this complex set of relationships for all sorts of reasons – the most obvious being that music is produced by human and instrumental movement, and is thus indelibly stamped with its bodily and instrumental origins."

(Clarke & Davidson, 1998:76)

So, a review of the empirical evidence on motion and e-motion based on the iconic dynamic extroversive coding, and having these two factors as directly connected with the role of the body in music performance, will allow for a better understanding of the referred complex set of relationships in the generation of music's expressiveness.

i) Motion as the intentional meaning of the expressive music/performance

In the later 1930s, Truslit (1938) empirically established a strong connection between music and motion, focusing on the information conveyed in the sound pattern. According to Shove & Repp (1995):

"He contended that the musical dynamics and agogics (timing variations) convey movement information directly to sensitive listener, who can then instantiate these movements by acting them out, if necessary. For Truslit, the goal of music performance is to arrange the musical surface in accord with the appropriate underlying movement."

(Shove & Repp, 1995:71)

Truslit identified three basic types of curves (open, closed and winding), and has considered that the essential thing in music performance is to experience the music as the expression of a natural motion according to a combination of them. He proposed that the combination of larger (more global) and smaller (more local) movements in the performance (as well as in the listener's physical response to them) reflects the organisation of the motor system into two main divisions: one controlling whole body movement (associated with the vestibular organ, which controls body balance, orientation, and our sense of self-motion) and a more peripheral one, which controls the movements of

the limbs (see Clarke, 2001:215). Work by Todd (1992, 1995, 1999) presents a similar speculation based on the physiology of the inner ear. He considers the hypothesis of the sound being able to activate the vestibular apparatus and thus the perception of motion in expressive musical performance to originate in an internal sense of motion. He argues for a distinction between two periodic kinematic components in human response to music – a vestibular component and a sensorimotor component. The former associated with small-scale gestural motion, the latter with large-scale whole body movements. In a close line of investigation to Todd, other researchers (Sundberg & Verrillo, 1980; Repp 1992, 1993 Kronman & Sundberg, 1987) also have shown that "timing patterns of spontaneous performance follow the curve of objects moving in a gravitational field (…), suggesting that what sounds natural in performance does so because it mimics the behaviour of physical objects moving in the real world" (Clarke & Davidson, 1998:77).

As Feldman et al. (1992) pointed out, a performer who produces tempo changes is not under a physical constraint to follow any particular function. The constraint to produce tempo or other musical parameters whatsoever depends first and foremost of mental, structural and expressive purposes as well as all the constraints derived from the fact that a performer has to interact not only with the instrument, but also with co-performers and audience, acoustic and physical space, social, and even historical conditions. However, the finding that a performer intending to create an expressive musical performance mimics aspects of human movements, including the movements characteristic of affective and emotional life, leads to the argument that motion is an important and fundamental part of the expressive intention of music performance communication process. Thus, the music performance is clearly a process of communication and interaction between performer(s) and audience while accomplishing artistic functions of musical expression and meaning. One of Sloboda's (Sloboda, 1998) key discussion points considers as a priority to understand dynamically musical meaning the fact that the human being occupies social, physical and biological worlds:

> "Dynamic awareness of music involves reading the music as an embodiment of something
> else, and my proposal for that something else is, broadly, the physical world in motion,
> including that very special sub-class of moving objects, the living organism."
>
> (Sloboda, 1998:28)

And, also, in other words:

> "It is the whole process of being a biological human inhabiting a physical and social world
> that provides the fuel and the impetus for a dynamic understanding of music."
>
> (Sloboda, 1998:25)

It seems thus that much would be gained in contemporary music if the dynamic (physical, biological and social) understanding and engagement with music would be seriously taken into account by composers and performers. Snook (1987), for instance, suggests that contemporary composers who are interested in creating works essentially connected with 'sound gestures' should concern themselves with creating a choreography of a performer's movements rather than attempting to write a score in the traditional sense. According to Shove & Repp (1995: 79):

> "Much music composed in this century encourages only primitive forms of motion or
> inhibits natural motion altogether Many twentieth-century composers focus on sound
> qualities or on abstract tonal patterns, and performers of their compositions often neglect
> whatever kinematic potential the music may have. The absence of natural motion
> information may be a significant factor limiting the appreciation of such music by
> audiences.
>
> (Shove & Repp, 1995:79)

Another author that clearly seems to corroborate this state of ideas is Meyer (2001). He claims that:

> "The empathy that belief begets is not only cognitive. Almost always it involves physical
> behaviour – an inner performative empathy, a kind of imitative identification, with the
> qualities and patterns of music. Lack of motor empathy, for example, seems to have played
> a significant role in the negative aesthetic and hence emotional response to avant-garde
> music, while motorically accessible musics (from Bach to rock) have found ardent
> audiences."
>
> (Meyer, 2001:353)

Cognitive perceptual engagement and physical emphatic identification with natural or biological sources of motion seems thus an indispensable part of musical meaning and, hence, of e-motion perception and response to music performance.

ii) E-motion as the intentional meaning of the expressive music/performance.

In a parallel but closely related line of investigation, the study of the e-motional impact of the expressive music performance started to show that whether the music is 'fully improvised' or 'structurally provided', as in a notated score, it always comes to the performance's core as essentially the same: *an act of intentional meaning embodied in musical experience* (Stubley, 1966; Clifton, 1983; Davidson & Correia, 2001; Cox, 2001).

Gabrielsson (1973b) noticed that the listener's sense of motion in rhythm might be connected to three main but different dimensions of the rhythmically perceived phenomena: a structural dimension ('experienced structure of the rhythms'), a motional dimension ('experience motion of the rhythms'), and an emotional dimension ('emotional aspects of the rhythms experienced'). Later, Gabrielsson (1982) provided evidence showing how rhythmical patterns that were identical in the notation were nevertheless performed differently in pieces of different emotional character. Later on, Gabrielsson (1995) investigated the effects of the performer's expressive intentions concerning the emotional character of the microstructure in music performance. He presented results which supported the theory of the existence of an "isomorphism between the structure of emotions and the structure of music, as proposed by Langer & Clynes (suggesting that the isomorphism should include motion as well)" (Gabrielsson, 1995:46). The isomorphism studied was with special regard to the microstructure in music performance, i.e., how musicians use factors such as deviations in timing, articulation, dynamic range, amplitude envelopes, vibrato, in order to generate various emotional characters. Nevertheless, Gabrielsson concluded that: "this isomorphism principle is applicable to all levels, from the shaping of the individual tone up to large musical forms" (Gabrielsson, 1995:46). Another author, Shaffer (1989) while discussing the possibility of a robot's construction in order to play a Chopin's waltz, realised the insufficiency of the motor programming and cognitive planning in accomplishing that task. He concluded that emotional factors must be considered as well in order to catch fully the mood of the music. Later, Shaffer (1995)

explicitly referred to the task of the performer (as he says, a 'privileged listener') as being "to interpret the structure of a piece of music and choose a character in terms of patterning of timing, dynamics, timbre and articulation that sympathetically conveys the structure" (Shaffer, 1995:18). In a former study, he argued that a performer can have freedom to shape its moods while remaining faithful to its structure:

> "Expression, including the choice of tempo, seems relevant to conveying mood as an aspect of musical meaning. If we further suppose that music can convey an abstract narrative, then we can think of the musical structure as describing an implicit event, and the gestures of musical expression as corresponding to the emotional gestures of an implicit protagonist who witnesses or participates in the event. Thus, the performer's interpretation can be viewed as helping to define the character of the protagonist..."

> (Shaffer, 1992:265)

Clarke (1995), referring to Shaffer's (1992) study, argues that he introduces the concept of 'character' or 'narrative' in relation to expressive music performance based in three counts:

> "First, it emphasises that expression cannot simply be mapped onto structure; secondly, it draws attention to the performer as a creator of musical character; and thirdly, it provides a link with the idea that performers unconsciously use physical gestures associated with emotional states as a basis for shaping musical expression."
> (Clarke, 1995:26)

In a former work, Shaffer (1984) had come to the conclusion that while some of the timing deviations from the metric pulse of the musical rhythm are the inevitable expression of features of the motor program, other timing profiles variations are flexible in tone production. Though integrated in the motor program these features can be seen as timing modifications intended to emphasise the musical effect. In line with Shaffer's evidence, Nakamura (1987) has demonstrated that differently intentioned performances contain different quantities of dynamic variation consciously used and recognised by performers as expressive devices. So, different performance parameters like timing, dynamics, intonation, and timbre may be manipulated according to an individual's interpretation. The work of Davidson (1994) was able to bring again some empirical evidence on the relationship between some of the parameters of sound expression in performance and the

performer's body movements. For instance, she was able to establish a strong connection between the intensity of music sound expression and the movement size – the more intense the expressive music intention, the larger the body movement of the performer. In fact, empirical research by Friberg & Sundberg (1999) also suggested that the body is able to provide a source for the generation of musical expression. But, interestingly, Davidson (1993, 1994) also noticed that besides these responsive to sounds body movement types, the performer could intentionally use other movements that seemed not to be so strictly bound to the musical structure. Davidson & Correia (2002) argue that perceivers can detect finely grained information about musical expression (timing, pitch, and dynamic modifications to structural features of the music) and intention (the emotional mood of the performer and the piece) from a performer's body movements. According to them, musical performance movements are of the following types and have clear functions and meanings:

> "Musical performance movements are of the following types: purely biomechanical (...), individual (...), and culturally determined (...). Many of these performance movements have clear functions and meanings: to communicate the expressive intention (...); to communicate directly with the audience or coperformers about issues of coordination or participation (...); to signal extramusical concerns (...); to present information about the performer's personality, with his or her individualised characteristics providing important cues (...); and to show off the audience."

> (Davidson & Correia, 2002:244-45)

It seems thus that the more recent work of Davidson adds some critical empirical evidence to the idea that the performer's intention might be also to emphasise important features of the musical structure that are not directly connected with stylistic and syntactical implications. This may be possible through the use of expressive movements that, although acknowledging the musical structure and remaining intrinsically connected to its presence, reach beyond its intrinsic stylistic and syntactical meaning. Or even, as Davidson points out, through expressive gestures that may exist independently of the structure but corroborating the intended meaning of it. These expressive gestures may be seen as contributing to the 'character' and 'tone' of the music's implied emotional meaning, as suggested by Hatten's explanation of Pierce's 'tone of voice' notion. It seems thus, that the best way to understand how the 'iconic-dynamic-extroversive coding' functions in music

performance is by establishing a parallel relationship with the expressive performance of human voice.

Historically, the idea that both proto-speech and proto-music have evolved from primitive affect vocalisations has been suggested by philosophers like Rousseau (1761) or natural scientists like Helmholtz (1863). Nowadays, following up on these suggestions, we may find reference in some authors (Sundberg, 1982; Vaneechoutte & Skoyles, 1988; Kivy, 1989; Scherer, 1991; Lavy, 2001; Juslin, 2001) arguing that music, like language, evolved out of primitive vocalisations, which have originally served to communicate basic emotional states, such as happiness or sadness. Scherer (1995:236), for instance, claims that "vocalisation, which remained a major modality for analogue emotion expression, became the production system for the highly formalised, segmental systems of language and singing". According to him:

> "In speech, changes in fundamental frequency (Fo), formant structure, or characteristics of the glottal source spectrum can, depending on the language and context, serve to communicate phonological contrasts, syntactic choices, pragmatic meaning, or emotional expression. Similarly, in music, melody, harmonic structure, or timing may reflect the composer's intentions, depending on specific traditions of music, and may simultaneously induce strong emotional moods. This fusion of two signal systems, which are quite different in function and structure, into a single underlying production mechanism, vocalisation, has proven to be singularly efficient for the purpose of communication."

(Scherer, 1995:236)

In fact, and despite our sophisticated level of linguistic communication, vocal communication is often made in a non-linguistic use of paralinguistic elements of the vocal utterance. The production of paralinguistic modulations involves physiological processes quite different from the linguistic sounds. Speech production involves highly cognitive processes controlled by areas of the neo-cortex (Damasio, 1994) while non-linguistic utterances are primarily controlled by the limbic system and emotion systems (Scherer, 1995:240). Over the last few decades, scientists have investigated how emotion might be inferred acoustically from different vocal utterances. Scherer (1995:240) argues that "given the manifold determinants of voice production processes, even slight changes in physiological regulation will produce variations in the acoustic patterns of the speech

waveform". Kappas, Hess, & Scherer (1991) provided a synthesis of research in the field, in which they point to the importance of acoustical parameters such as intensity, intonation, and fundamental frequency in determining emotional state. Anger, for instance showed an increase in fundamental frequency (F_0) and intensity with downward-directed F_0 contours; sadness, presented lower fundamental frequency and low intensity; fear, high intensity and increase in fundamental frequency. The research based on the expression of emotions by actors' voices offered a way to examine the timbral quality of the acoustic signal through the upper partials present in the signal: anger, for instance, showed very strong upper partials while sadness evoked almost none. Fonágy (1976) pointed out that different emotions have different articulatory effects, with anger showing tense movements and extreme articulatory positions, while tenderness being associated with slow and gentle patterns of movement. Wagner (1989) suggests that much of the variation in fundamental phonation frequency, timbral quality, tempo and articulation occurring in the acoustic signal of emotional voices could be attributed to the effect of ANS arousal in the vocal system. Lavy (2001:43), based on a plethora of empirical studies showing a high degree of consistency between the listener's judgements and a wide range of emotional expressions, states that "emotion encoded in an acoustic signal really is communicated". He also argues that this may happen because communication and recognition of emotion encoded in acoustic signals, such as voice, forms part of an innate and universal human signalling system that operates independently of culture-specific knowledge. In fact, Frick (1985), in his cross-cultural study of vocalised emotions, has demonstrated that many emotions and their vocalisations exist and are recognised universally. Development psychology studies (Papousek, 1996; Papaeliou & Trevarthen, 1998) have shown that the use of vocal utterances, known as 'motherese', to express and communicate emotional states is a universal and innate trace of human kind. According to Scherer (1992) this could be based on an apparently innate human propensity to mimic:

> "The automatic tendency toward motor mimicry or observed expressive behaviour serves as the mechanism for emotion transfer. In other words, the observation of someone else's expressed emotion will, at least in a rudimentary fashion, evoke the same emotion in us, the observer."
>
> (Scherer, 1992:55-56)

In the same way, Scherer (1992) speculates, we may perceive a vocal signal as a physical expression and through an automatic process of motor mimicry we might be able to feel that very same expression. That is what Walton calls the 'imitative covert behaviour' (Walton, 1997). In fact, empirical evidence from studies of motor imagery (Gallese & Goldmann, 1998; Fadiga & Gallese, 1997; Fadiga et al., 1998) report that understanding the observed behaviour of others involves imagining performing the same or similar actions. Thus, it might be concluded, together with Baddeley & Logie's (1992) argument, that if comprehension of spoken words ought involves covert imitation it seems reasonable to expect that comprehension of sung words ought to involve covert imitation as well. Lavy's (2001:48) motto 'music as utterance' explores the possibility that an accurate account of the musical expression of emotion might be based upon the human voice as expressive of emotion:

> "If the perception of emotive cues encoded within vocalisation is a universal phenomenon, and the acoustic profiles of those cues is consistent pan-culturally, we should expect that any musical sound that shares an acoustic profile with such an utterance also has the potential to carry the relevant emotional sign. We should expect further that signs communicated by such musical sounds would be perceived and could be responded to in a way that is no different from perception of and response to any other utterance; in other words, emotional response to such music would be, at least in part, emotional response to the apparent human utterance. In many cultures and societies – including our own – one of music's primary roles has been to communicate emotion; it is here claimed that it owes much of its success to the fact that such music either consists literally of vocal expression, or else emulates vocal expression by means of analogous sounds. The parameters along which musical sounds very are the same as those along which emotional cues in vocalisations vary; the building blocks of music are those vary same units that form the basis of non-verbal communication."

(Lavy, 2001:48)

It has already been stated (see Chapter 2, pg. 86) that in the Western music tradition, composers' intentions have been often focusing in the communication of emotion. Bel Canto composers have set in music all kind of emotional repertoires and have tried to express through the voice of the singers the feelings and the emotions of human kind. Caccini's (1614) declaration in his introduction to *Le Nuove Musiche* claimed, as referred,

for a music that should honour the natural contours of vocal expression. The sounds used in music by those composers are thus formalisations of the sounds that are called for by the expressed emotions when expressed as a vocalisation or as a human utterance, and the kind and degree of formalisation is dictated according to the particular stylistic system within which, and according to the instruments for which, the music is composed. Meyer (1956:260), by noticing the implicit connection between musical patterns and vocalised emotional expressions, stated that "because moods and sentiments attain their most precise articulation through vocal inflection, it is possible for music to imitate the sounds of emotional behaviour with some precision." Thus, it is possible to say that 'music can *sound like* the vocal expression of emotion' (Budd, 1985:132). Striking examples of what has just been stated can be found in some cross-cultural musical genres like, for instance, the lullaby. Trehub et al. (1993), for instance, carried out a cross-cultural study on this musical form, and found that lullabies are clearly recognised and differentiated from non-lullabies all over the world, regardless of the musical culture from which they originated. Lullabies are sung in a quiet, gentle voice, using patterns (repetitive motifs, slow moving contours) that typify the sound of human vocal utterances that express relaxation and contentment. Given the clear function of lullabies and the patterns used, it seems possible to conclude that they have clear similarities with the related soothing vocal sounds.

Research on the perception of emotional expression in musical performance has brought empirical evidence that demonstrates that pure instrumental music produce sounds that might be heard as human utterance and presents similar acoustical cues to the sounds produced by the human voice when expressing an emotional state. Gabrielsson & Juslin (1996:85) showed that regardless of instrument, there was a core set of music expressive features that typified various emotions: 'anger', was normally expressed by quick tempo, high intensity, marked contrasts between long and short notes, harsh timbre and distortion, sharp tone onsets, mostly non-legato articulation; by contrast, 'sadness' was typified by slow tempi, relatively large deviations in timing, a low intensity sound lacking in high harmonics and a slow vibrato, legato articulation, slow tone onsets; 'fear' was characterised by irregularity of time, very large deviations in timing, large dynamic variation, mostly staccato articulation, very fast vibrato and variations in sound intensity; 'happiness' was expressed by bright timbre, fast mean tempo, moderate timing variability,

staccato articulation, fast and light vibrato, sharp contrasts between long and short notes, moderate to loud sound level, rapid tone onsets. The importance of this particular research has been that though Gabrielsson and Juslin have worked with instrumental musicians, yet they have found patterns of results that match the speech and song patterns of emotional communication. This seems to suggest that acoustical signals involved in the expression of emotion in instrumental music are the same as those involved in vocalisation and, therefore, in vocal music. The fact that a musical sound that carries emotional cues behaves and is perceived as if it were a human voice communicating emotional meaning makes arguable the perspective that suggests that the production and perception of emotional states through vocalisation is a human cross-cultural and innate ability. Besides that, it tells us that this cross-cultural and innate ability has an equivalence in the human capacity to perceive such emotional signals in a form of an acoustic waveform and therefore respond to them in a way that is differentiated from the human response to other types of sounds. Music performers that intend the communication of emotional meaning tend to produce sounds with similar acoustic profiles – timbral and dynamic features – to those of emotive vocal utterances. Ultimately, it seems arguable that the expressive gestures of such performers match in quality and dynamic the expressive intentional gestures of emotive vocal utterances. This seems to be in conformity to what Clynes (1969) meant when he explained that phonatory and articulatory gestures are manifestations of a common expressive dynamic form, which underlies both the perception and production of expression in different modalities. A plethora of empirical research and evidence on the production and perception of emotional acoustical cues have been undertaken since then. Much of this more up-to-date empirical evidence corroborates and complements the early findings of the very first research on the subject (Hevner, 1936; Gundlach, 1935; Rigg, 1939; Watson, 1942).

Among the factors that affect emotional expression in music structure, 'tempo' (presto, allegro, moderato, andante, adagio) seems to be usually considered the most important (Gundlach 1935; Hevner, 1937; Rigg, 1964; Scherer & Oshinsky, 1977; Juslin, 1997). Fast tempo for happiness, anger, fear, surprise, excitement and potency among others; and slow tempo for sadness, tenderness, disgust and calmness among others. Nevertheless, structural factors should be studied together and in musical context, because, for instance, although

faster tempi may point to the expression of the emotion happiness, other factors may overrule this tendency (minor mode, chromatic scales or sharp tone onsets) and thus influence the final result. 'Mode' is the second major important factor that affects emotional expression in music structure: major mode (Hevner, 1936; Rigg, 1939; Kleinen, 1968; Wedin, 1972; Scherer & Oshinsky, 1977; Crowder, 1985; and Krumhansl, 1997) is usually determinant of happiness/joy as perceived emotion. Minor mode, instead, is mainly implicative of sadness as the type of expression perceived (Hevner, 1936; Rigg, 1939; Kleinen, 1968; Wedin, 1972; Crowder, 1985; and Krumhansl, 1997). Of course, which expression is perceived depends, again, on the general music context. Other major factors are 'pitch', 'loudness', 'harmony', 'tonality', 'rhythm', 'articulation', 'melody', 'intervals', 'musical form', etc. Here are some of the different implications that these factors may generate in the perception of emotional expressions. High-pitched intervals can be indicative of happiness and low-pitched intervals of sadness (Maher, 1980). Loud music may be associated with intensity/power (Wedin, 1972; Kleinen, 1968), anger (Juslin, 1997) and joy (Rigg, 1939) and soft music with tenderness (Kleinen, 1968; Juslin, 1997) or sadness (Juslin, 1997). High pitch may be connected to happiness (Hevner, 1936) or anger, fear and surprise (Scherer & Oshinsky, 1977), and low pitch to sadness (Hevner, 1937; Watson, 1942; Wedin, 1972; Scherer & Oshinsky, 1977) or solemnity (Rigg, 1940; Wedin, 1972), for instance. Simple, consonant harmony may be indicative of happiness (Hevner, 1937; Watson, 1942), gaiety and pleasantness (Wedin, 1972) and harmony complex/dissonant implicative of, for instance, sadness (Hevner, 1937; Watson, 1942), anger (Lindström, 1997) and fear (Krumhansl, 1997). Regular/smooth rhythm may be perceived as expressing, for instance, happiness (Watson, 1942), varied rhythm may be perceived as expressing joy (Thompson & Robitaille, 1992), flowing/fluent rhythm as expressing gaiety (Wedin, 1972), complex rhythm as expressing anger (Thompson & Robitaille, 1992), and firm rhythm as expressing sadness (Wedin, 1972). Staccato articulation is mostly implicative of gaiety (Nielzen & Cesarec, 1982) or fear and anger (Juslin, 1997) and legato articulation of softness (Wedin, 1972) or of tenderness and sadness (Juslin, 1997). (A detailed review on earlier and later findings on the interface between musical structure and emotional expression can be found in Gabrielsson & Lindström, 2001).

The review just presented has thus focused upon those factors of the composed musical structure that are represented in the musical notation and contribute to the perception of emotional expression (tempo, loudness, pitch, mode, melody, rhythm, harmony and other formal properties). Therefore, a music performer should be aware of these factors in order to be able to understand and underline, through his/her articulatory gestures, the emotional meaning already presented in the musical notation.

A complementary but slightly different approach to the emotional expression in music attempts to find out which emotions can be reliably expressed in music performance, and by which processes and features the emotion is communicated in music performance. [42] According to Juslin (2001), performance is an important element in determining the emotional expression. Early studies of emotional communication in music performance (Kotlyar & Morozov, 1976; Senju & Ohgushi, 1987) were already concerned with investigating whether performers were able to communicate emotions to their listeners. Different researchers have tried to describe the means by which performers express specific emotions (Baroni & Finarelli, 1994; Gabrielsson & Lindström, 1995; Gabrielsson & Juslin, 1996; Rapoport, 1996; Juslin, 1997, 2000; Juslin & Medison, 1999). One main finding from this line of research is that "the performer's expressive intention affects almost every aspect of the performance: that is, emotional expression in performance seems to involve a whole set of *cues* – or bits of information - that are used by performers and listeners" (Juslin, 2001:314). Thus, Juslin's opinion is that to be able to study the communication of emotion in music performance as a process, we must consider the 'code' used by performers and listeners. This code is made up of the set and interplay of the performance expressive cues, including the use of tempo, sound level, timing, intonation, articulation, timbre, tone attacks, tone decays, and pauses. Mean level of a cue and its variability throughout the performance seems to be important for the process of communication. Sadness, for instance, is usually communicated through the use of a slow tempo, low sound level, legato articulation, small articulation variability, slow tone attacks, and dull timbre, whereas happiness can be communicated by fast tempo, high sound level,

---

[42] The latter approach has been the focus of a detailed review by Juslin, 2001. Scherer & Zentner (2001) present another complementary approach for the study and classification of emotion carried by music. They propose for future research "adopting a systematic approach for deriving a taxonomy of musical emotions while focusing on the aroused, not the perceived, emotions" (Scherer & Zentner, 2001:381).

staccato articulation, large articulation variability, fast tone attacks, and bright timbre (see Juslin, 2001). Juslin shares, as referred, with other authors (Helmholtz, 1954; Kivy, 1989; Scherer, 1991; Sundberg, 1982) the idea that there is a close relationship between music and the human voice. He suggests that the hypothesis that there is an iconic similarity between vocal expression of emotion and musical expression of emotion applies mainly "to those aspects of the music that the performer can control during his or her performance, such as tempo, loudness, or timbre" (Juslin, 2001:321). [43] Juslin (1995, 1997) has elaborated a theoretical framework aimed specifically at performance aspects, which adopts a functionalist perspective on emotional communication in music performance (see Chapter 2, pg. 88). This framework involves the integration of ideas from research on emotion and non-verbal communication with Brunswik's (1956) probabilistic functionalism. According to the functionalist perspective, the fact that performers are able to communicate emotions to listeners might be explained by considering the 'functions' that communication of emotion has served in general social interaction (Darwin, 1872). Juslin argues that although music performance may be a recent addition to the human behavioural repertoire, it must still be constrained by psychological mechanisms that have been shaped by evolution (Juslin, 1998). He suggests that the origin of the code used by performers and listeners reflect two main factors: brain programmes and social learning. Juslin (2001), claims the existence of innate 'brain programmes' for vocal expression of emotion is supported by evolutionary studies (Plutchik, 1994) and by neuropsychological studies (Jürgens & von Cramon, 1982; Panksepp, 1982; Ploog, 1986) which present evidence of 'brain programmes' that function to initiate and organise 'pre-wired' vocal expressions. [44] Thus, the functionalist perspective searches to specify the actual neurological substrates that are responsible for the similarities in the communication of emotion between vocal expression and music performance and undertakes empirical research that may provide evidence in favour of this hypothesis. Having this in mind,

---

[43] Neuropsychological evidence suggests that certain aspects of music, such as timbre, share the same neural resources as speech, whereas others (e.g. tonality) draw on neural resources that are unique to music (Handel, 1991; Patel & Peretz, 1997). Further, the same brain hemisphere (the right) seems to be involved in the perception of emotion in music and speech (Borod et al., 1998; Bryden et al., 1982).

[44] Juslin (2001:321) explains that, according to the supra-cited neuropsychological studies, "electrical stimulation of particular brain structures in the amygdala and the hypothalamus produces vocalisations that observers cannot distinguish from natural expressions of emotions. In humans, at least, these 'pre-wired' vocal expressions can also be executed at will for communicative purposes, thanks to the anterior cingulate cortex."

Juslin (2000, 2001) carried out a systematic review of the code usage in vocal expression and music performance. The results provided clear evidence of a number of similarities in code usage. For example, vocal expression of sadness is associated with slow speech rate, low voice intensity, low intonation, and little high-frequency energy in the spectrum of the voice (Scherer, 1986, 1995; Banse & Scherer, 1996). Empirical evidence from music performance research (Gabrielsson, 1995; Juslin, 2000; Scherer, 1995; Scherer & Siegwart, 1995) confirms that the same acoustic cues are used to communicate sadness in musical performances. Juslin (2001) adds that, besides the innate brain programmes for vocal expression, certain aspects of emotional expression in music performance may be governed by social learning and cultural influences. This modulation of the code might be seen as starting with the early interaction between mother and infant (Papousek, 1996; Trevarthen, 1999). Sloboda (1996a) has shown that performers make links between expressive cues and extra-musical aspects through analogies. Teachers, performance conventions, literature and critics certainly help to modulate expressive intentions throughout a performer's carrier. Nevertheless, Juslin (2001) decidedly assumes that the core of the expressive code is provided by the first factor – the innate brain programmes.

In fact, Adachi & Trehub (1998) have shown empirical evidence suggesting that children between 4 and 12 years old seem to be able to use some of the innate brain programmes for vocal expression based cues to communicate and express emotions in song. For instance, they use a fast tempo and a high sound level to express happiness and a slow tempo and low sound level to express sadness. Singing is clearly at the intersection of vocal and music performance expression, and therefore it has often be seen as a model to investigate the ensemble of acoustic cues that may constitute the emotional expressive code in music performance (Seashore, 1937; Fairbanks, 1940; Kotlyar & Morozov, 1976; Födermayr & Deutsch, 1994; Gabrielsson, 1994; Scherer & Siegwart, 1995; Gabrielsson & Juslin, 1996; Rapoport, 1996; Sundberg, 1999).

For instance, Sundberg, Iwarsson & Hagegard (1995) claim that pitch characteristics, particularly overtone content, are especially important emotional attributes in singing. In their study of a professional baritone singing short phrases mostly 'Lieder', highly experienced singing teachers were asked to say in what ways the second of two

performances (the first being neutral and the second with appropriate expression) differed in relation to seven standard emotions: secure, loving, sad, happy, scared, angry and hateful. Despite the difficulty shown by the results to define expression in single adjective terms, a clear division appeared corroborating the intentional expressiveness of the second performance. Sundberg and co-workers analysed the different recordings in terms of tempo, vowel-to-vowel duration, overall and short-term variation in dynamic, frequency oscillation and formant frequencies. Results in tempo and vowel to vowel variation were identified according to Todd's principle of expressive timing. The largest source of variability in sound levels was linked to emphatic pronunciation of consonants. In agitated pieces vibrato was greater in expressive performances than in deadpan versions, while in non-agitated music less vibrato was used in expressive performances than in deadpan versions. Analysis of formant frequencies showed 'colouring' of voice timbre in expressive versions. Sundberg et al.'s (1995:229) conclusions suggest that the use of timing variations are clearly linked to structural features but other expressive cues are instead linked to characteristics of speech studies and vocal expression of emotion.

Rapoport (1996) systematically classified eight categories of timbral structures in singing and correlated them with specific emotional expressions. The timbral structures were analysed from single tones by Fast Fourier Transform, and range from neutral/soft to expressive, excited and virtuoso. Rapoport (1996:131) claims that there is an exact correlation between the shape of a singing signal in a FFT spectogram and the vocal folds tension. Based on Clyne's (1973, 1977) 'sentic' theory, which claims the existence of biologically pre-programmed spatiotemporal forms, or *essentic forms*, for the communication of specific emotions, it seems possible to argue for a clear relationship between the state of vocal musculature and the tension and relaxation associated with various states of emotional arousal and expression. Rapoport (1996) has demonstrated that the very first unit pulse of any tone provides already enough information about the emotional character of the singing performance. Rapoport's (1996:140) analysis of two different performances of "Morro, ma prima in grazia" from 'Un Ballo di Maschera' by Verdi has revealed very different timbral structures. According to him, this difference does not originate in musical structure or, even, text. He explains that Margret Price's performance reflects 'deep sorrow' and 'acceptance' of her tragic fate, while Maria

Callas's performance expresses more 'pathos' and even 'protest' because she does not accept that she has to die.

Scherer & Siegwart (1995) carried out a similar study on the importance of emotional feeling and emotional communication on the opera stage. Two excerpts from the cadenza in "Ardi gli incensi" from Donizetti's opera 'Lucia de Lammermoor' were acoustically analysed for five different performances of the cadenza by Toti dal Monte, Maria Callas, Renata Scotto, Joan Sutherland, and Edita Gruberova. The acoustic parameters of the singing voices were correlated with preference and emotional expression judgements, based on pairwise comparisons, made by a group of experienced listener-judges. In the different analysis of the acoustic parameters, they tried to find out which vocal cues affected the most the listener's judgements. Scherer & Siegwart (1995:259) presented the following conclusions: a) the different interpretations elicited significantly different listener ratings of emotional expressiveness; b) the voice samples of the five singers differ quite substantially with respect to objective acoustic variables; and c) it is possible successfully to predict listener attributions on the basis of the objective acoustic characteristics (for further detail see Scherer & Siegwart, 1995). Focusing again on the extroversive dynamic coding and the expression of emotion in music performance, the dynamic variables of singing in the time domain (connected or not with linguistic and semantic content) such as tempo, pausing, vowel duration, phrasing, tone transitions, and others, like pitch contour and timbre are also highly variable and determined in emotionally expressive singing. Scherer & Siegwart's (1995) study focused exclusively on acoustic measurements of voice quality, in the sense of spectral rather than temporal characteristics of the singing voice. Sundberg et al. (1995), as mentioned before, focused more on different recordings in terms of tempo, vowel-to-vowel duration, overall and short-term variation in dynamic, frequency oscillation and formant frequencies. Therefore, the study of different variables will help to elucidate on how and by which means emotion can be expressed and perceived in music performance communication.

To conclude several results of a comparison study by Gabrielsson (1995) will be presented in order to elucidate how each of these vocal emotional expression connected variables can be present and recognisable in the performance of music with different instruments and be

affected in specific ways for each emotional expression. Gabrielsson (1995) asked musicians to play the synthesiser, violin, flute, electric guitar and a singer in order to perform short pieces of music, some well-known, others less well-known or unknown, with the intention of having them sound happy, sad, solemn, angry, soft/tender, and without expression. They were asked to keep the pitches in the melodies but were free to use any other means to achieve the required expression (see Gabrielsson, 1995). Digital analysis of the recorded performances showed, for instance, that:

> "Tempo was usually more rapid for happy and angry versions than for the others; sound level was loudest for angry versions and softest for tender or sad versions; dotted patterns were performed sharp in angry versions but soft in sad and tender versions; articulation was 'airy' in happy versions, legato or legatissimo in sad and tender versions; 'pitch bending' was used in sad and tender versions; tone onsets and offsets were abrupt in angry versions but slowly rising/decaying in sad and tender versions; timbre was manipulated in different ways, for example, 'harsh' in angry versions, 'cold' in inexpressive versions; vibrato was deleted in inexpressive versions."

> (Gabrielsson, 1999:549)

All the empirical evidence reviewed here suggests and corroborates the idea of a main corpus of emotion expressiveness in music/performance directly or indirectly connected to the articulatory and interactive movements of the performer, and directly connected to certain specific dynamic characteristics and patterns of the musical structure.

THE ICONIC-STRUCTURAL/DYNAMIC CODING FUNCTIONING

To conclude, these two proposed levels of emotional coding in music performance, iconic-dynamic-introversive and iconic-dynamic-extroversive coding seem to be, in fact, sub-levels of the same case of music's emotional character coding – the *iconic structural dynamic coding*. These sub-levels of the 'iconic structural dynamic coding' are actually both part of the relatively tight iconic intersection between the musical structural features and musical dynamic implications, the performer's articulatory movements, and the processes of subjectivity that they call for, i.e., music patterns and their dynamic

implications, patterns of movement whose general characteristics are similar to bodily movements symptomatic of human emotions, and internal states somatically experienced.

To the *Iconic-structural/dynamic coding* belongs all the empirical research concerned with the expressive musical event whose conception of music/performance is based on a awareness of the dynamic implications of the performed music's structure, as well as of the dynamic and structure of the performer's articulatory movements. In ecological terms, it corresponds to the core of the *expressive musical event*. The perception of motion that characterises it involves an extraction of the relational invariants, which specify structure and transformation of the performer and constitute the structure and transformation of an imagined 'musical object'. Embodied in the forms of music, these relational invariants specific to natural classes of movement include (in the present coding) those movements symptomatic of human emotional states, i.e., 'movement similar to bodily movement symptomatic of human emotions' (Shove & Repp, 1995:61). Thus, according to the iconic structural dynamic level of coding of the musical expressive event, and from an ecological perspective, the relational invariants extracted by the perceiver from the musical performance event will help him/her to structure his/her awareness of it in terms of recognisable e-motional patterns, such as those characteristics, for instance, of happiness or sadness.

But, while in the *introversive coding* sub-level the embodied emotional character is, as said, an achievement of the dynamic sensations of flux, tension, implications and expectations fulfilled or violated that are latent in the musical structure and that can be emphasised or frustrated, directly or indirectly, through the articulatory movements of the performer, in the *extroversive coding* sub-level the music emotional character is an achievement of the relationships that come about through some formal resemblance between the musical structure and some event or agent carrying emotional 'tone'. Thus, in the latter case, the expressive event remains directly or indirectly connected to the articulatory and interactive movements of the performer and directly connected to certain specific dynamic characteristics and patterns of the musical structure, conveyed in the performed music, and able to generate emotion through an iconic relationship with the inherent structural features of the emotional states elicited.

D. Fourth, the *Symbolic-associative coding*, through which the expressive event might remain indirectly connected to the articulatory movements of the performer, and to all the factors of the performed music that can be connected with emotional content in an individual's or collective's memory, often due to occasionally and idiosyncratically learned associations and conditioning (see Scherer & Zentner (2001), and Sloboda (2001).[45]

---

[45] Though music, through its structural features, offers the possibility of investing and of being invested with meaning, this investment is not constituted through an univocal relationship between music and meaning. It is not, as it is in language, an arbitrary and conventional relationship between structured sounds and processes of subjectivity. The symbolic meaning is based, as said, on a purely arbitrary and conventional relationship between the sign and its object, to be understood only by virtue of a system of signs of which it is a part. This one-to-one relationship, as in the relationship between a word and its meaning, does not exist in music at the level of the expressive musical event. At least, it does not seem appropriate that this 'one-to-one relationship' should be considered as a part of *the Iconic-structural/dynamic coding*. On the contrary, it seems wiser to avoid this conception specially in the case of the *iconic-dynamic-introversive coding* where by reason of the implicative processes of structural and stylistic devices of music/performance and their associated meanings - based on a "process of evolved, sociocultural conventionalisation" (Scherer & Zentner, 2001:364) - one might be tempted to consider it as a part of a symbolic-associative way of meaning. This quasi-arbitrary almost 'one-to-one' relationship way of conveying emotional information is almost never to be found in music. If it exists, it is more appropriate to think it as a part of a different level of awareness of the musical event. This conception has been considered in connection to what has been here called the *Symbolic-associative coding*. It has been suggested (see *associative coding* in Scherer & Zentner, 2001:364), that in this case no necessary relationship between the inherent characteristics of the sound and the inherent characteristics of the phenomenon evoked or referred seems to be required. Often, it has been stated that it is a rather common difficulty to present a hard-and-fast distinction between iconic and symbolic functions in musical meaning (Clarke, 1995). Lavy (2001: 9) pointed out the difficulty to place Meyer's work, and the empirical evidence that supports it, between icon and symbol. Nevertheless, and according to Shepherd & Wicke (1997:178), the relationship between specific structural elements (music patterns and their implications, realisations or non-realisations) of the performed music and the processes of subjectivity that they call forth in individuals, is to be better understood as an 'iconic intersection', because it is based on a necessary relationship between the inherent characteristics of the music and the inherent characteristics of the phenomenon elicited. Thus, in the case of examples such as those found in Cooke's (1959) theory, it might be said that the proposed one-to-one relationship between sounds and meanings seems to fit better the *symbolic-associative coding* (as proposed here) then the *iconic-dynamic coding* type. Nevertheless, we still may wonder if the associative process between the musical elements and their socio-culturally conventionalised meanings had any sort of a 'structural-trigger' that might 'justify' historically the one-to-one relationship proposed. Perhaps, the one-to-one relationship between the music sounds and the suggested emotional meanings, in Cooke's (1959) theory, is indeed a result of a cultural 'conventionalisation' through a constant repetition and usage of the predictably triggered meanings of these sounds. Davies (1994:172), for instance, argues that, "though Cooke believes that composers do express their feelings, he does allow that expressiveness is inherent to the character of musical sounds." Lavy (2001:49) suggests that Cooke's (1959) examples of Emotion-related musical patterns might be abstractions and formalisations of natural vocal expression. Shepherd & Wicke (1997) seem to share a similar opinion: "It would thus seem inappropriate to think of these kinds of sounds in music as giving rise to signification in a *purely* arbitrary or conventional fashion. Such sounds may well have become conventionalised through constant and repeated usage in terms of the predictable meanings to which they are *likely* to give rise (...). However, it should not be forgotten that a structural basis for the meanings constructed through this dimension of sound in music has to lie behind their conventionalised aspects" (Shepherd & Wicke, 1997:158).

The 'emotion-triggers' of the *symbolic-associative coding* might be seen as conventionalised musical meanings that are most of the time just the product of contingent, local, and arbitrary associations. In other words, between the inherent characteristics of the sound and the inherent characteristics of the phenomenon evoked or referred, the relationship found is often of an idiosyncratic type. The fact that they are idiosyncratic does not mean necessarily that these associations are not capable of generating a strong emotional experience. In fact, Gabrielsson (1991, 2001) and Sloboda (1991) have found examples of specific pieces of music that in this sense are strong emotion's triggers. Sloboda (2001) claims that such emotions tend to lead attention away from the present music onto the remembered past event. In some cases, common cultural experiences may lead to shared emotions that are essentially the result of an associative process. According to Sloboda (2001:95), it is the case of the extreme negative emotions triggered by Wagner's music in many Jews after the Second World War. These triggered emotional meanings are, according to him, based in associative sources of emotion, and do not require any particular feature of the musical structure for their elicitation (see discussion in footnote 45). Nevertheless, some of the performer's articulatory movements may be seen as making part and influencing the symbolic-associative coding. For instance, Davidson & Correia (2002:244) suggest that some of the musical performance movements may be learned through imitation of others' behaviours in specific cultural contexts and so may have common presentations and perhaps emotional meanings, independent of their binding to any particular feature of the musical structure.

3.4 SUMMARY

It was stated that recent investigations into music/performance report that a significant part of its creativity is connected to the performer's capacity to express and communicate emotion (see, for a review Sloboda, 1996b and Sloboda & Juslin, 2001). Thus, whether the music is 'fully improvised' or 'structurally provided', as in a notated score, it always comes to the core of the performance as essentially the same: *an act of intentional meaning embodied in musical experience* (Stubley, 1966; Clifton, 1983; Davidson & Correia, 2001; Cox, 2001). Thus, when music is presented within a performance neither is its structure absent from the presention meaning nor is its meaning 'emptied' through the presence of the musical structure. Of course, in case of the expressive musical event, meaning implies

emotional meaning. Music is, therefore, a structured and structuring medium for *purposes* which reach beyond structure but whose existence remains intrinsically connected to its presence. It has long been stated, but it seems nowadays better understood and confirmed by empirical evidence, that there is an homology of inherent organisation and dynamics between the sounds of music, the movements and dynamics of our affective life, and the patterns of movement 'whose general characteristics are similar to bodily movement symptomatic of human emotions, moods or feelings' (Shove & Repp, 1995:58). It has been suggested that a semiological framework might aid to recognise and organise the communication process of the emotional character of music/performance:

- First, a significant part of the *expressive musical event* is directly connected to the articulatory movements of the performer, and to the performer's capacity to express and communicate emotion within the bodily movements and gestures ["Indexical gestures" (Hatten, 1999: CSI Lecture II, 4); "Illustrative and emblematic gestures expressing" (Davidson, 2002:242-3)] that articulate the performed music and interact with other performance elements (instruments, co-performers and audience) critical in the generation of emotional meaning within the musical performance. And, also, directly connected to the acoustic structure of small segments of sounds (timbres, for instance) ["Segmental features" (Scherer, 2001:362)] and others aspects of the sound quality of the performed music (causally related, for instance, to the physical effort required to the production of the sound). (Indexical-qualitative coding).

- Second, a less large but not less significant part of the *expressive event* remains directly or indirectly connected to the articulatory and interactive movements of the performer ["Spanning & Arcing" (Hatten/Pierce, 1999, CSI Lecture III, 7-8); "Structural-connected-expressive gestures" (Davidson, 1991); "Analogical Mapping" (Sloboda, 1998:25)], and is a direct implication of specific and conventionalised structural and transformational features ('symptomatic' of human emotional states) ["Iconic Introversive/Supra-segmental Features-Iconic/Symbolic Coding" (Scherer, 2001:364) and "Native and Syntactic Processes" (Meyer, 2001: 344) and "Intrinsic sources of emotions" (Sloboda & Juslin, 2001)] of the performed music. (Iconic-structural/dynamic-introversive coding).

- Third, another significant part of the *expressive musical event* remains directly or indirectly connected to the articulatory and interactive movements of the performer ["Tone of voice" (Hatten/Pierce, 1999 CSI Lecture III, 7-8); "Specific-expressive gestures" (Davidson, 2002:242-3); "Expressive Cues/Countours" (Juslin, 2001: 319)], and directly connected to certain specific dynamic characteristics and patterns of the musical structure ["Iconic Extroversive/ Supra-segmental Features-Iconic/Symbolic Coding" (Scherer, 2001:362) and "Statistical Parameters" (Meyer, 2001:342) and 'Native processes" (Meyer, 2001:344); Extrinsic-Iconic sources of emotions (Sloboda & Juslin, 2001)] existent in the performed music and able to elicit emotion through structural resemblance. (Iconic-structural/dynamic-extroversive coding).

- Fourth, a larger but less significant part of the expressive event remains indirectly connected to the articulatory and interactive movements of the performer ["Type-token gestures" (Hatten/Pierce, 1999 CSI Lecture II, 6); "Adaptative gestures" (Davidson, 2002:242-3); Imitation gestures (Davidson & Correia, 2002:242-3)], and indirectly to certain patterns and specific dynamic characteristics of the musical structure and to the structural and transformational features of the performed music ["Associative Coding" Scherer, 2001: 364); "Associative sources of emotions" (Sloboda & Juslin, 2001)]. All these factors 'can be associated with emotional content in an individual's or collective's memory due to learned associations and conditioning' (Scherer & Zentner, 2001:364), often, occasionally and idiosyncratically generated. It means, there is no necessary relationship between the inherent characteristics of the sound and the inherent characteristics of the phenomenon evoked or referred to, only an arbitrary and contingent one. It is, as Davies (1978) put it, the typical case of *'Darling, they're playing our tune'* theories. (Symbolic-associative coding).

Having now developed a theoretical framework, which provides a structure for how emotion and musical meaning interface, it is important to explore the extent to which these ideas can be investigated empirically. Indeed whilst theoretical propositions are key ways to engage with understanding, it seems that there is a need for empirical evidence if these theories are to be used in Teaching and performance practice. So, it is upon this theoretical ground that the next chapter emerges.

CHAPTER 4

4. EMPIRICAL EVIDENCE ON SINGING AND EMOTION

EMPIRICAL STUDIES ON THE EXPRESSION OF E-MOTION IN THE FACE AND VOICE OF THE SINGER IN THE COMMUNICATION OF EMOTIONAL MEANING IN SINGING

4.1 INTRODUCTION

Having examined in Chapter-3, from a semiotic point of view, the different ways by which the 'emotion character' of music/performance may be conveyed according to different levels and sub-levels of coding, emotional meaning will be now empirically investigated within the context of singing.

The empirical investigation that follows will focus on the expression and perception of emotion examining the specific acoustical (voice) and visual (face) features of the emotional communication process in singing. The process of this investigation will make use of the different explained levels and sub-levels of coding, either separately or/and combined, in order to be able to understand and explain the communication of emotion in the performance of singing. The complex interface of all these different levels of the music emotional communication coding process can be seen at work in the analysis of the song *Die Post* presented in Appendix F. However, and for the sake of the framing of the experimental studies in the scope of this chapter, it should be added that the more direct *indexical-qualitative coding* elements of the music performance emotional communication are the main focus of this investigation. Thus, the empirical investigation undertaken is mainly concerned with the facial and vocal expressive elements of the music performance emotional communication which are also a part of the emotional expression in everyday communication. Nevertheless, as is discussed in Appendix F, some of the emotional expressive vocal and facial elements may also be enhancing some of the music structural features identified in Chapter-3 as they can also belong to the Iconic-structural/dynamic (introversive/extroversive) coding. For example, when voice and face emotional expressiveness enhance the emotional meaning of the music structural features in bars 12 to 16 in the song *Die Post*, in Experiment-IV (Iconic-structural-extroversive coding); or when in the same experiment the vocal and facial emotional expressive elements enhance the emotional content of the music structural features in bars 37 to 46 (Iconic-structural-introversive coding). Empirical evidence has revealed that the expression of emotion in a

musical performance may arouse emotions in listeners (Hatfield et al., 1994). Hatfield et al. (1994:5) have called 'emotional contagion' to the "tendency to automatically mimic and synchronise facial expressions, vocalisations, postures, and movements with those of another person and, consequently, to converge emotionally." Also, Sloboda's investigation (1991, 1992) has found empirical evidence of specific structural features associated with the elicitation of bodily and behavioural manifestations of emotion (tears, thrills, etc). Sloboda et al. (1997) have found empirical evidence supporting the fact that performers "may enhance listeners' emotional responses to the music by emphasising notes that are of particular 'significance' in the composition, thereby enhancing violations of musical expectations that are already latent in the structure. Otherwise, Juslin (2001:330) states that "a performer can also arouse emotions in listeners by performing in a manner that deviates from stylistic expectations with regard to the performance of certain structure." Thus, it seems possible to state that expressive performances might induce emotional responses in listeners both ways, i.e., introversive and/or extroversive structural dynamic coding. In fact, it may well be that some of the recognition process of emotional perception in music performance is in fact concomitant with the listener's response to induced or aroused emotion (see Krumhansl, 1997). This will surely be a further step of future research on emotional meaning in singing. Nevertheless, because the main focus of this thesis is the study of expression and perception of emotion in singing, and because it is also a part of its aim to elaborate a 'cognitive feedback tool' to assist singers and singing teachers to check and improve the expressiveness of their performances, the empirical approach of this investigation has not considered to take in account the arousal perspective on emotional meaning.

4.2 METHODOLOGICAL QUESTIONS

The general basic questions to be asked by the empirical investigation of this thesis were: how to relate expression in singing (acoustically and visually) to perception of emotional content; how to make singers and singing students aware of their expressive intentions and at the same time aware of the achievements of their trials when singing with emotional meaning; and how to give them appropriate ecologically valid cognitive feedback of the expressive cues used when performing a vocal piece with intentional emotional meaning.

4.2.1 CATEGORICAL VS. DIMENSIONAL APPROACH TO EMOTIONAL EXPRESSION

The different ways people express, recognise and experience emotions is certainly of capital importance in the manner how subjects differentiate emotions and emotional contents within artistic communication. The two main approaches to understand and classify emotional behaviour are the dimensional approach and the categorical approach.

The dimensional approach

The dimensional approach focuses on the identification of emotions based on their placement on a two, or three, dimensional structure (valence, activity, and potency). Russell's (1980) dimensional approach is a circumplex model of emotion - a two dimensional, circular structure, based on the bi-polar dimensions of valence (like *vs* dislike) and arousal (high *vs* low). Analysis of judgements of emotion words or facial expressions using factor analysis or multidimensional scaling is the most common of this kind of approach. The purpose of Russell's (1980) model and other models or formulations (Wundt, 1987; Woodworth, 1938; Schlosberg, 1941) seem, in fact, more oriented to investigate music mood-induction than perceived emotional categories. In fact, dimensional models seem especially accurate to capture the continuous changing of emotional expression occurring, for instance, during a piece of music (Sloboda, 2001:78). As Scherer & Zentner (2001) claim, emotion theorists proposing dimensional models generally focus on the subjective feeling rather than the expressive components of emotion (Scherer & Zentner, 2001:376). The reason lies in the fact that the bi-polar dimensions of valence and arousal (and potency, see Osgood et al. (1957) explanation for a third proposed dimension interpreted in terms of the semantic dimensions of the linguistic signs) are simultaneously part of the content of different categorical perceived emotions, and they are not sufficiently clear in themselves to make an emotion recognisable. This fact, however, seems to be of crucial relevance to the perception of the performed vocal music, according to Salgado (2000, 2001, 2002a, 2002b). The interpretation of emotional meaning in the performance of vocal music seems to be highly dependent on the performer's ability to express distinctly, i.e., categorically, the emotional content of music either vocally as facially (see Experiment III, pg. 171). Thus, the expression and perception of emotional meaning in singing seems directly connected to the performer's capacity to distinctly express the emotional meaning of the music.

The categorical approach

According to the categorical approach, listeners experience emotions as categories that are distinct from each other. The essential aspect for the defenders of the categorical approach (see Tomkins, 1962; Izard, 1977; Ekman, 1992; Oatley, 1992; Plutchnik, 1994) is the concept that basic emotions are of a limited number, innate and universal, and all the other emotional states derive from them. According to Scherer (1995:235), "the most important feature of the emotion mechanism is that it produces specific action readiness while providing a latency period that allows adaptation of the behavioural reactions to the situational demands." So, they can be seen as adaptive in life emergency situations (ready for execution) but also as way of externalising a reaction to a specific emergent situation, and a way of communicating this information to the social environment. According to Darwin's (1872) pioneering work on the expression of emotion in human and animals, emotion is phylogenetically continuous and is based on complex and interactive relationships among the different social elements. According to Scherer (1995:235), "all expressive modalities, particularly body posture, facial features, and vocalisation, are involved in emotion communication." And, as far as facial expression is concerned, he adds that, "behavioural and social scientists have made much progress in gathering evidence on the phylogenetic continuity, the universality across cultures, and the rich information content of emotion in facial expressions" (Scherer, 1995:235). In this sense, Oatley (1992) suggested that each basic emotion might be defined in evolutionary terms by the functional adaptive strategies dependent on the evaluation of relevant goal events. Some of these relationships between the functional adaptive strategies and the evaluation of relevant goal events are what Figure III shows in an adapted version from Sloboda & Juslin's (2001:76) table 4.1:

Emotion	Juncture of plan (a)	Core relational theme (b)
Happiness	Sub-goals being achieved	Making a reasonable progress towards a goal
Anger	Active plan frustrated	A demeaning offence against me and mine
Sadness	Failure of major plan or loss of active goal	Having experienced an irrevocable loss.
Fear	Self-preservation goal threatened or goal conflict	Facing an immediate, concrete, or overwhelming physical danger

Figure III - Key appraisals for basic emotions according to Oatley (a), 1992 and Lazarus (b), 1991. (Adapted from Sloboda & Juslin, 2001:76).

The innate and phylogenetic continuity character of basic emotional expressions seems to be supported by Izard (1971), whose definition of facial expression points "not merely to an expressive behaviour but, rather, to a pattern of neuro-muscular activity that constitutes a component of emotion." Spitz (1963:60) claims also that, though the development of emotional expressions occurs via "object relations", they develop from an innate "Anlage" (pre-disposition). The earliest systematic experimentation on muscle movements was done by Duchenne (1862). He found out that the use of small electric currents to the sites of the different facial muscles, produce independent movement of these muscles. Based on Darwin's (1872) facial expressions landmark book (his idea was that certain facial expressions were biologically produced, and that these will be universally recognisable), Duchenne's (1862) physician work studies, and Hjortsjö's (1969) work concerning the description of the visible appearance changes for each muscle of the facial expressive movements, Ekman and Friesen (1978) elaborated a FACS manual (Facial Action Coding System) in order to describe in a comprehensive and detailed way the different facial actions. [46] With this system they were able to codify different facial action units, and provided a reliable way to describe and understand the facial actions that make part of the different emotional facial expressions. Ekman (1989) also provided empirical evidence for the universality of basic emotion's facial expressions. Modern investigation (Calder et al., 1996), based on computer-graphics-techniques, has recently provided a relevant way that seems to corroborate empirically Ekman's claim on experiencing emotions categorically. Testing the fundamental issue concerning whether we recognise facial expressions categorically or dimensionally, Calder et. al. (1996) studied the perception of photographic-quality 'morphed' images of facial expressions. Their conclusions seem to be consonant with categorical perception in other domains like colour perception, which means that the perceived emotion changes abruptly when the perceiver comes across different categories.

---

[46] Facial Action Coding (FACS), is an anatomically based method of measurement which distinguishes 44 different facial action units, allowing an accurate description of facial configurations, and can specify the point at which each movement begins and ends. It is a descriptive tool used to assess whether a particular expression corresponds to an emotion's prototypical facial expression (Ekman, 1993).
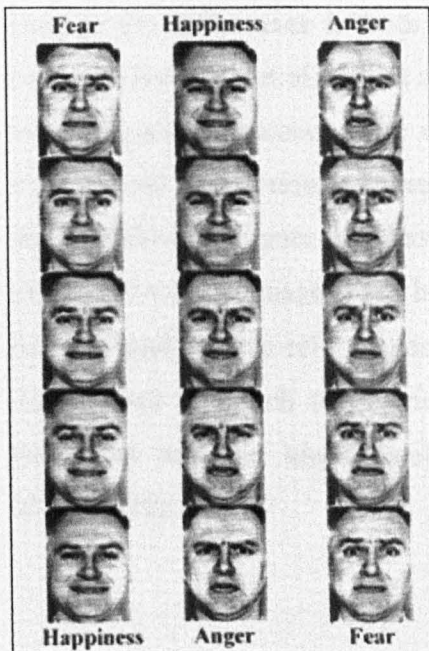
Figure IV - . Shows three of the facial expressions of basic emotions from the Ekman & Friesen (1976) series (Happiness, Fear, and Anger) used by Calder et al. (1996) to create morphed images of facial expressions for continua which run from fear to happiness (first column), happiness to anger (second column), and anger to fear (third column). In each column, the images are moved 10%, 30%, 50%, 70%, and 90% of the distance from one expression to the other. People were asked whether the expression on each face was most like fear, happiness, or anger. Identification rates suggested that there were abrupt shifts between perceiving one emotion and the other near the mid-point of each continuum. Adapted from Bruce and Young (1998:197).

On the other hand, Frick (1985) and Scherer (1995) claimed, based on listener's ability to correctly recognise vocal emotional expressions, that the vocal expression of at least some emotions (a list where all the basic emotions are included but also love, pride, jealousy) seems to be universal. Scherer (1981) presented a review of different studies that appeared over the last 50 years on the recognition of emotional states from voice samples and reported an average accuracy of ~ 60%. Bezooijen (1984), for instance, reported a vocal emotion recognition mean accuracy of 65% for emotions such as, disgust, surprise, shame, interest, joy, fear, sadness and anger. Later, Scherer et al. (1991) reported for fear, joy, sadness, anger and disgust, a mean accuracy of 56%. Salgado (2000) reported a mean accuracy for basic emotions' vocal expression in singing of ~ 65%.

A general criterion has been suggested for distinguishing basic emotions from 'secondary' or 'complex' ones. According to Sloboda and Juslin (2001:77), the categorical approach has formulated that, basic emotions:

> "(a) have distinct functions that contribute to individual survival; (b) are found in all cultures; (c) are experienced as unique feeling states; (d) appear early in development; (e) are associated with distinct patterns of physiological changes; (f) can be inferred in other primates; and (g) have distinct emotional expressions."
>
> (Sloboda & Juslin, 2001:77)

Despite the differences in both approaches, some recent contributions to the study of emotion (Nyklicek et al., 1997) seem, nowadays, more inclined to regard the dimensional and categorical approach as a useful and complementary tool for the investigation of emotion and of emotional expression. The prototype approach to emotion (Rosch, 1978), shown below in Figure V, "based on the idea that language and knowledge structures associated with language shape how people conceptualise and categorise information", can be seen (and is often referred) as an "interesting compromise between the categorical and dimensional approach to emotion because it addresses both the content of individual categories and the hierarchical relationships among categories" (Sloboda & Juslin, 2001:78-79).
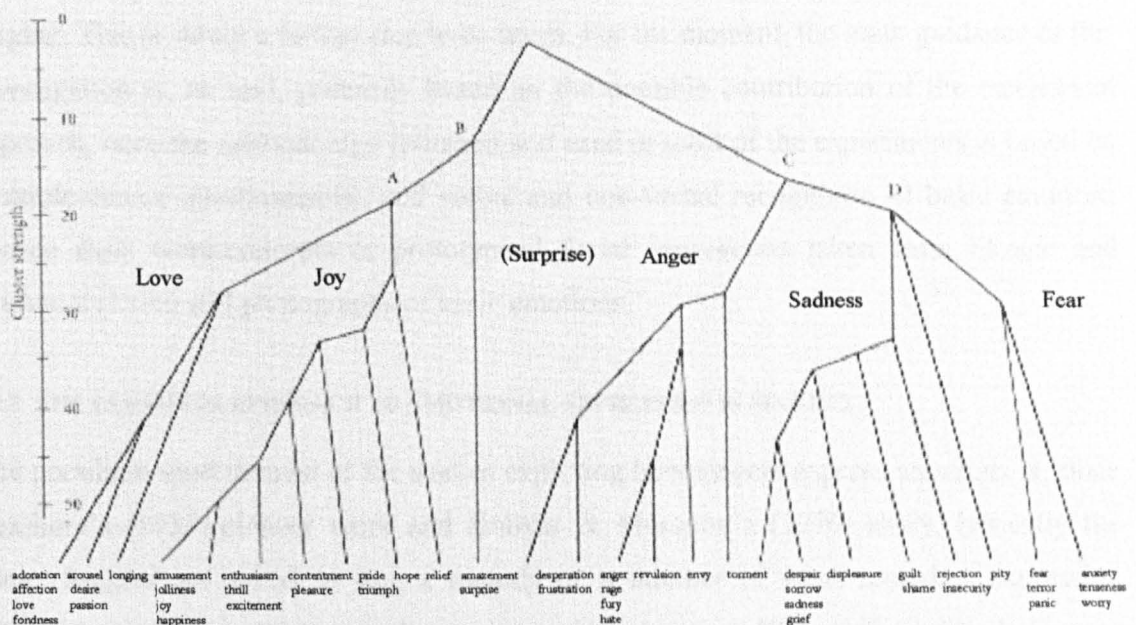
Figure V - . A prototype analysis of emotions words (from Shaver et al. 1987). The vertical dimension of the structure shows the hierarchical relationships among categories. The most general level is the *superordinate* level, which is defined by the positive or negative valence of the emotions within a particular category. The middle level represents the basic-level categories, or *prototypes*, of emotions, which anchor our mental representations of all emotions within a given category. The *subordinate* level consists of all other emotions related to a particular prototype. Finally, the horizontal dimension shows the relations among members of the same prototype category (see Shaver et. al., 1987). Adapted from Sloboda & Juslin (2001:80).

Of course, some space should be also left to the non-verbal and non-linguistic approach to the study of emotion (and thus to the study of emotional meaning communication in music

performance), once it seems that the structures subjacent to the expression, recognition and communication of emotion are not all dependent on the structures of language, and some of them are already a part of the human behaviour pattern even before the existence of the structured language. Indeed, according to Izard (1971), and following Lorenz's (1965) conception of *behaviour patterns*, it seems that "emotions and their expressions are strongly influenced by the information encoded in the genome" (Izard, 1971:58). Given the assumptions that the face is "the primary theatre of emotions" (Ekman, 1992), and that either dimensional and categorical approaches take facial expressions as a powerful source of emotional data, and that the performance of Lied (song) is basically expressive at the level of the face (facial expressions, movements and articulation), it would be expected to find much and interesting evidence for complementary work of these different investigation approaches applied to music performance, in general, and particularly to singing. This is surely a further step to be taken. For the moment, the main guidance of this investigation is, as said, primarily based on the possible contribution of the categorical approach, once the methodology followed and used in most of the experiments is based on multiple-choice questionnaires, and verbal and non-verbal recognition of basic emotions chosen from word-concepts or prototypical facial expressions taken from Ekman and Friesen's (1976) still photographs of basic emotions.

4.2.2 THE PARADIGM APPROACH TO EMOTIONAL EXPRESSION IN SINGING

The paradigm used in most of the studies exploring how singers express emotions is, since Seashore's (1937) pioneer work and Kotlyar & Morozov's (1976) study, basically the same. Singers are asked to sing a melody or a number of short melodies expressing different emotions according to the researcher's instruction. The performed melodies are firstly recorded or videotaped and then listened to, watched and evaluated in order to check whether listeners are able to recognise the intended expression. Each performance may then be further analysed in order to study what acoustic or visual means the singer has used to achieve each emotional expression Though there is substantial empirical evidence on the acoustic means used by singers the investigation of the visual means, to date, is practically unexplored. Hence, it is an important aim of the current investigation.

The use of real music excerpts to study emotional expression in singing ensures good ecological validity (see Gabrielsson & Lindström, 2001 for details of naturalistic enquiry and different methodological approaches to music and emotion). Also, the common singing of the same melody with different kinds of emotional expression secures the internal validity of the experiment, because the recognition of the conveyed emotion should be mainly the result of the singer's expressive intentions (a device used by Davidson (1993) in the exploration of expressive intention in music performance). All these experiments have focused, for the reasons exposed, on the perceived emotion rather than on the induced emotion. The emotions under investigation have usually been sadness, happiness, fear, anger, and tenderness. The studies have mainly focus on two things: the accuracy of the communication and the code used by the singers and the listeners (see Juslin, 2001) to re-present the emotion. According to Gabrielsson & Lidström (2001), listeners usually reported the perceived expression by either:

> "(a) free phenomenological descriptions; (b) choice among descriptive terms, adjectives, or nouns, provided by the investigator; (c) ratings of how well such descriptive terms applied to the music in question. Free descriptions were subjected to content analysis. Listener's choice among descriptive terms was analysed regarding the frequency with which each term was chosen and regarding inter-subject agreement. Ratings were usually analysed by multivariate techniques, such as factor analysis, cluster analysis, and multidimensional scaling in order to find a limited number of fundamental descriptive dimensions. Recently, various techniques for continuous and/or non-verbal recording of perceived expression have been developed."

> (Gabrielsson & Lindström, 2001:224)

According to these assumptions, the series of experiments in this thesis was undertaken in order to study the communication of emotion in singing. In these experiments, two different angles were adopted in order to elaborate an expressive analytical tool that could provide singers, singing teachers and students of singing with a reliable cognitive feed-back about their capacities to communicate emotional meaning within singing:

- To investigate the production of the facial movements and vocal sounds directly involved within the expression of emotion in singing;

- To explore how the performed elements are perceived, recognised and experienced by the audience.

The examples of the singing repertoire chosen for the empirical investigation have been all chosen from the Lied repertoire because the main focus of this study is connected to the facial and vocal expression in singers and the interpretative characteristics of the Lied repertoire oblige the singer to express the musical meaning almost exclusively through these two main channels of communication.

4.3 EXPERIMENTAL STUDIES ON EMOTIONAL MEANING IN SINGING

INVESTIGATING THE PRODUCTION OF FACIAL MOVEMENTS AND VOCAL SOUNDS INVOLVED IN THE COMMUNICATION OF EMOTION IN SINGING

Four different experiments are reported and discussed in this Chapter. They all draw on the same performers and the general procedure for the production of the sung stimuli was identical in all the three first experiments, though they differ in the procedure for the perception of the stimuli. The fourth experiment, though drawing on the same performers, and on the same number of perceivers, had a totally different procedure for both the production and perception of the sung stimuli. The shared features are described below, but in order to contextualise the purposes of the four experiments, their aims are described now:

Experiment I, aimed to study the emotion facial expression in the performance of singing in a quantitative-objective way. To do this, the singer's facial gestures and how they were connected to the expression of emotional meaning in singing was investigated. This was done using a purely quantitative measurement without reference to the subjectivism of the audience's perception.

Experiment II, aimed to study the emotion facial expression in the performance of singing in an objective way, through the perception and recognition of point-light displays of the different emotional expressions performed. Thus, how and if a certain audience perceives and recognises the emotional meaning of the dynamic information contained in the abstract movement of the reflective markers (point-lights) resulting from the filmed singer's different expressions of emotional meaning in Experiment-I was investigated.

Experiment III, aimed to study the perception and recognition of emotional meaning within the performance of singing, through the observation of videotaped performances of singers singing the same musical phrase with different emotional contents. Thus, it was investigated empirically and objectively how a certain audience perceives and recognises the emotions communicated through the voice and the face of the singers within different videotaped performances of singing.

Experiment IV, aimed to study the perception and recognition of emotional meaning within the performance of singing in a real concert situation. That is, how a certain audience perceives and recognises the emotional meaning communicated through the voice and face of the singer when performing.

General methodology for experiments I, II and III
Participants
Two experienced professional singers, one female (singer 1) and one male (singer 2), with an average age of 40 years old, were the participants for the empirical observation. These individuals will be referred throughout by their names: Silvia and Toni.

Audience
Fifteen participants acted as observers, with an average age of 25-30 years. A mixed audience, with equal number of men and women, representing a combination of musicians, amateur listeners and less experienced observers.

General procedure for the stimuli production
Silvia and Toni were asked to prepare and sing a phrase of the Lied repertoire and express different emotional states in different performances (happiness, sadness, anger, fear, and neutral). The chosen musical phrase was an extract from Schubert's Erlkönig song, the 3 bars between bar 123 and bar 125 of the song (see Figure VI). This musical phrase was chosen both for being an emotional peak within the musical structure of the Lied, and also because of the neutrality of the sung text, i.e., the words could apply to almost any situation or emotional state: "Mein Vater, mein Vater!" (My Father, my Father!). Note that Fig. VI shows the original version of the Schubert 3 bars, however each singer chose a pitch range to suit their own voice, so Silvis began on A3 and Toni on B2.

Figure VI - Shows the 3 bars between bar 123 and bar 125 of the original version (Ed. Peters, pg. 175) from Schubert's 'Erlkönig' song. The phrase 'Mein Vater, Mein Vater' was sung in five different emotional conditions: anger, fear, happiness and sadness, and neutral.

The musical line itself both rises and falls within a limited range of a minor second, and so does not make particular technical demands on the singer, and so again leaves interpretative possibilities open. Recordings were made in five different conditions: neutral, happy, sad, fearful, and angry. These four fundamental emotions were chosen because they are well-reported as being the most strong and clearly recognisable in many contexts, even in cross-cultural contexts (Izard, 1980, 1994), and belonging to the set of basic emotions recognised as innate and universal (Ekman, 1989, 1992, 1994; Izard, 1994). Neutral was used as a control measure, as a sort of baseline from which the singer was then to "emote" upon. It was of added interest to see if the audience could differentiate between the neutral condition and the emotional states. Each emotional state was sung fourteen times by each singer. So, all together, seventy different excerpts formed the raw material used for the first three experiments. From these, five renditions by each singer were selected among the most coherent and representative of the emotions performed (the selection was made according to the singer's self evaluation, together with the research team's suggestion, of the most expressive and authentic versions of the several attempts for each emotional state). Clarke & Davidson (1998) used multiple performances of the same piece in their investigation of a pianist's interpretation. They found a high level of interpretative coherence across performances. Of the six repetitions initially recorded two were selected as being representative of the whole data pool for a very detailed analysis. On these grounds it was decided to ask the singers to perform in the region of 12-15 replications so that the five most coherent and representative of the emotions could then be selected for detailed analysis.

As the data were being collected, the two singers decided to undertake 14 repetitions. These sung stimuli were filmed using a ProReflex camera system, tracked and stored in a computer (Experiment-I and II), and also videotaped (Sony DSR 300 P-DVCAM video-camera) and recorded through a Betacam sound channel (Experiment-III). For Experiment-I and II a point-light technique was used according to Bassilli's (1978, 1979) experimental work. In his work on the perception of emotion in the human face, Bassilli (1978) provided empirical evidence to support the claim that even finer moving stimulus of inner intentions are perceptually meaningful. His studies show that the dynamic information from points of light on the face is sufficient to identify different expressions: happiness, sadness, surprise,

interest, disgust, fear, and anger. By comparing the results of ordinary displays with the results of point-light displays of these emotions, he provided evidence that both were produced with equal accuracy. This indicated, thus, that dynamic information was a principal perceptual source of emotion perception. By using four subjects, he was able to conclude that the movement patterns for each emotion were similar between subjects. Bassilli's work argues a good case for the ecological account of social perception in that it demonstrates that even the subtlest of kinematic information can convey complex information in a direct manner. Davidson (1991) also used the point-light technique to investigate the perception of expressive movement in music performance.

4.3.1 EXPERIMENT - I: *A QUANTITATIVE-OBJECTIVE STUDY OF EMOTIONAL FACIAL EXPRESSION IN THE PERFORMANCE OF SINGING, USING POINT-LIGHT TECHNIQUE.*

Aim

The aim of the current study was, as said, to investigate empirically how the facial gestures directly connected with the expression of emotional meaning in singing and thus if they could be captured using a purely quantitative measurement approach. Based on the stimuli produced through the fourteen attempts for each emotion sung by Toni, Experiment-I investigated empirically if and how the singer's facial expressions of emotional meaning might be detected through purely quantitative parameters without reference to the subjectivism of the audience's perception.

The underlying questions of the investigation in Experiment-I were:

- Is it possible to differentiate each of the 5 proposed conditions (4 basic emotions + 1 neutral or no-emotion) each expression in itself, and from one another, without the reference of the subjective act of perceiving and recognizing the expressed emotion?

- Can one compare the gestures of these 5 different emotional facial expressions without considering the emotional content attached to them?

In this sense, Experiment-I (as well as Experiment-II) might be seen as an important complement to Experiment–III. In fact, the aim of Experiment–I was to verify in an objective manner (by measuring the facial movements through the detected positions of reflective markers attached to the face of the singer) if there are quantitative differences contributing to the discreteness of the different facial expressions for each performed emotion. And, the aim of Experiment-II was to investigate if the kinematic information contained in the abstract movements of the filmed reflective markers (point-lights) on the singer's face was able to convey, in a clear and direct manner, the different intentions of the emotional meaning expressed. So, based on Bassilli (1978, 1979) and Davidson's (1991) experimental work, the 14 interpretations sung by Toni were measured through the use of reflective markers attached to the his face, as shown in Figure VII.
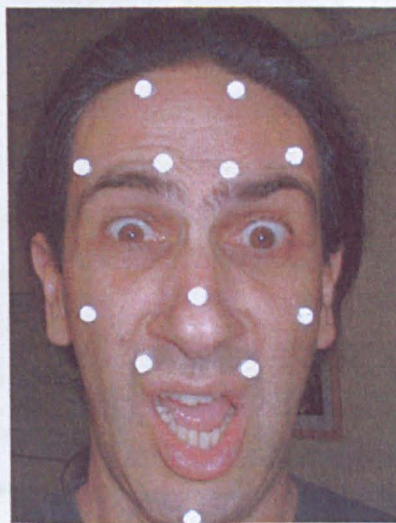
Figure VII - Markers on the singer's face

Method

In Experiment-I, the different performances of the song's phrase were recorded the ProReflex camera system, which captured the movement and the gestures of the face by detecting the position of reflective markers, attached, as said, to the head and face of Toni (see Figure VII). The exact location of the markers for the point-lights were chosen according to strategic points of the face where the muscles and the different muscles actions supposedly involved in the facial expression of the different emotions under study were detected in similar studies of everyday world emotional facial expressions. The referred studies are: Ekman & Friesen's (1978) - Facial Action Coding System, (FACS);

Izard's (1971), facial emotions photographs and charts description of the facial features, movements, and muscles involved in each of the basic emotions; and Hjortsjö's (1969), manual for the study of the effects of contractions of various facial muscles in everyday facial expressiveness (see Figure VIII).

Figure VIII - Some examples from Hjortsjö's (1969) manual for the study of the effects of contractions of various facial muscles in everyday facial expressiveness. 1. Lifting of *Frontalis (pars medialis)*, muscle facial movement typical of the facial expression of emotion fear; 2. Contracting of the muscles *Procerus, Levator Labii Superioris*, and *Alaeque nasi* typical of the facial of the emotion anger; 3. Contracting of the muscles *Zygomaticus major* and *Risorius* typical of the facial of the emotion happiness; 4. Contracting of the muscles *Triangularis* and *Depressor Labii Inferioris* typical of the facial of the emotion sadness. Adapted from Bruce & Young (1998:189).

The chosen places where the markers have been placed on the singer's face are referred as FR&L (Forehead Right and Left); BER&L (External Brow Right and Left); BIR&L (Inner Brow Right and Left); CL&R (Cheek Right and Left); LR&L (Lips Right & Left); and CHIN (Chin). A last marker was placed on the singer's bridge nose, a fixed point where no movement could be produced. So, the movement of all the other markers could be measured in relation to this fixed point, and that is why the distances of the several marker's movements are reported as a vertical distance (vdist). That is, the distance from the nose. By using simultaneously three different cameras to view the participant from different angles, the marker's position in space could be calculated from the "flat" pictures seen by each camera. The motion capture unit, ProReflex, with an appropriate algorithm, processed then each picture in real-time and calculated the coordinate of each marker in the scene. This information (raw 3-dimensional data) was then relayed to the computer, where the recorded movements and gestures were automatically calculated and stored by using the ProReflex software from Qualisys (run through a standard Windows 95 PC or Macintosh), which combined these features, made the set-up, calibration and post-processing. The difference between emotions was then quantified by comparing the average distance for each marker across the different facial expressions for each intended emotional musical meaning performed. This scheme was developed based on work done by Alan Wing at the Sensory Motor Neuroscience (SyMoN), Behavioural Brain Sciences

Centre of the University of Birmingham, U.K. These data were collected and stored in the computer's hardware of the laboratory and so are not available for illustrate purposes here (still images showing framed moments of the point-light movements in various emotional conditions can be seen in CD 1).

Analysis of data resulting from the measurement of the average distances across the different five emotions:

As explained, considering the large amount of data to be analysed, it was decided to consider for analysis only the best five of Toni's 14 attempted recordings. This choice was made according to the singer's self-assessment of the most authentic versions of the 14 attempts for each emotional state. Thus, for Happiness the most authentic versions chosen were HAP4, HAP7, HAP8, HAP10, HAP12; for Sadness were SAD14, SAD 16, SAD18, SAD23, SAD25: for Anger were ANG26, ANG27, ANG28, AND31, ANG33; for Fear were FEAR48, FEAR50, FEAR52, FEAR54, FEAR56; and for Neutral were NEUT38, NEUT40, NEUT42, NEUT44, NEUT46. These five attempts for each emotion were then played back, tracked and analyzed. The considered emotional conditions for this study were, as said, five – Happiness, Sadness, Anger, Fear, and Neutral – and the difference between emotions was quantified by comparing the average distance for each marker across the different expressions of the emotional meaning performed. The results of the measurements were presented in millimeters, and they showed the relevant differences between the distances of the exact same strategic face-points when measured across the different emotional expressions.

Table I shows the vertical distances for each marker through the 5 different chosen trials at the 12 strategically chosen face-points for each emotional meaning performed and also the average distance for each marker through the five different conditions (HAPav, NEUTav, SADav, ANGav, FEARav) at the 12 chosen face points for each emotional expression performed:

Table I: Vdist for each marker trough the five different conditions

TRIAL	FLvdist	FRvdist	BELvdist	BILvdist	BIRvdist	BERvdist	Nvdist	CLvdist	CRvdist	LLvdist	LRvdist	CHINvdist
HAP4	1.46	1.57	3.52	5.46	3.96	3.18	0.00	3.62	3.87	2.23	2.63	6.53
HAP7	1.43	1.36	2.88	4.79	3.78	3.25	0.00	3.96	4.17	2.49	2.57	6.94
HAP8	1.10	1.28	2.89	4.50	3.32	2.48	0.00	4.15	3.73	2.63	2.21	7.39
HAP10	1.50	1.38	3.37	5.33	3.90	3.26	0.00	4.23	3.63	3.67	3.79	9.03
HAP12	1.59	1.49	2.55	4.55	3.23	2.03	0.00	3.97	3.31	1.84	1.46	7.23
HAPav	**1.41**	**1.42**	**3.04**	**4.92**	**3.64**	**2.84**	**0.00**	**3.99**	**3.74**	**2.57**	**2.53**	**7.42**
NEUT38	1.61	1.34	1.82	1.80	1.37	1.31	0.00	2.08	1.56	0.99	1.62	6.01
NEUT40	0.95	0.74	0.95	0.68	0.53	0.51	0.00	1.19	0.98	0.62	0.61	4.08
NEUT42	0.66	0.46	0.67	0.47	0.31	0.32	0.00	0.99	0.94	0.47	0.64	4.44
NEUT44	0.59	0.53	0.56	0.37	0.37	0.46	0.00	1.20	1.19	0.56	0.75	4.66
NEUT46	0.40	0.42	0.40	0.25	0.29	0.46	0.00	1.25	1.26	0.50	0.72	5.28
NEUTav	**0.84**	**0.70**	**0.88**	**0.72**	**0.58**	**0.61**	**0.00**	**1.34**	**1.19**	**0.63**	**0.87**	**4.89**
SAD14	0.53	0.76	3.56	3.89	2.36	2.75	0.00	5.11	3.17	2.32	1.04	3.96
SAD16	0.61	0.59	2.69	2.53	11.81	1.95	0.00	4.14	5.76	1.59	0.77	4.00
SAD18	1.34	1.05	2.68	1.73	1.02	1.74	0.00	3.75	3.03	2.80	2.25	4.03
SAD23	0.79	0.50	2.60	2.87	1.37	1.99	0.00	5.32	4.65	2.86	1.83	3.80
SAD25	0.82	1.17	2.58	2.72	1.85	2.07	0.00	6.58	4.07	2.16	1.53	3.30
SADav	**0.82**	**0.81**	**2.82**	**2.75**	**3.68**	**2.10**	**0.00**	**4.98**	**4.13**	**2.34**	**1.49**	**3.82**
ANG26	2.39	1.92	4.45	4.27	2.56	3.88	0.00	4.37	4.33	3.44	4.43	10.48
ANG27	1.50	1.45	3.92	4.13	2.40	3.27	0.00	2.88	3.98	2.24	2.06	13.07
ANG28	1.94	1.44	2.86	3.06	1.43	2.04	0.00	2.25	17.13	1.87	2.12	8.65
ANG31	1.68	1.46	4.85	4.13	2.25	3.03	0.00	4.45	4.05	3.45	2.95	16.08
ANG33	1.47	1.30	4.40	4.81	3.96	4.21	0.00	2.77	3.07	2.81	2.69	16.73
ANGav	**1.80**	**1.51**	**4.10**	**4.08**	**2.52**	**3.29**	**0.00**	**3.34**	**6.51**	**2.76**	**2.85**	**13.00**
FEAR48	3.38	3.21	2.96	4.33	3.33	2.73	0.00	2.79	2.35	2.41	2.16	9.33
FEAR50	3.31	3.31	2.21	3.24	1.73	1.44	0.00	3.47	3.32	1.90	2.59	10.63
FEAR52	4.34	4.54	3.22	3.39	1.39	1.93	0.00	2.97	2.73	2.46	3.16	10.46
FEAR54	4.32	4.37	1.05	2.56	1.29	1.10	0.00	2.34	2.56	3.40	3.48	10.93
FEAR56	2.88	2.91	2.34	1.75	0.93	1.30	0.00	1.93	1.87	1.29	1.65	7.20
FEARav	**3.65**	**3.67**	**2.36**	**3.05**	**1.73**	**1.70**	**0.00**	**2.70**	**2.57**	**2.29**	**2.61**	**9.71**

Next, the average of the five trial's distances for each emotional condition was found for the full-face movements when performing the five emotional expressions (see Table II):

Table II: plotting for the full face movements

COND	FL&Rvdist	BEL&Rvdist	BIL&Rvdist	CL&Rvdist	LL&Rvdist	CHINvdist
	Forehead R & L	External brow R & L	Inner brow R & L	Cheek R & L	Lips R & L	Chin
HAPav	1.42	2.94	4.28	3.87	2.55	7.42
NEUTav	0.77	0.75	0.65	1.26	0.75	4.89
SADav	0.82	2.46	3.22	4.56	1.91	3.82
ANGav	1.66	3.69	3.30	4.93	2.81	13.00
FEARav	3.66	2.03	2.39	2.64	2.45	9.71

A graph was then printed out, showing (see Figure IX) the difference between the distances of the movements of the full-face emotional expressions (see explanation in discussion):
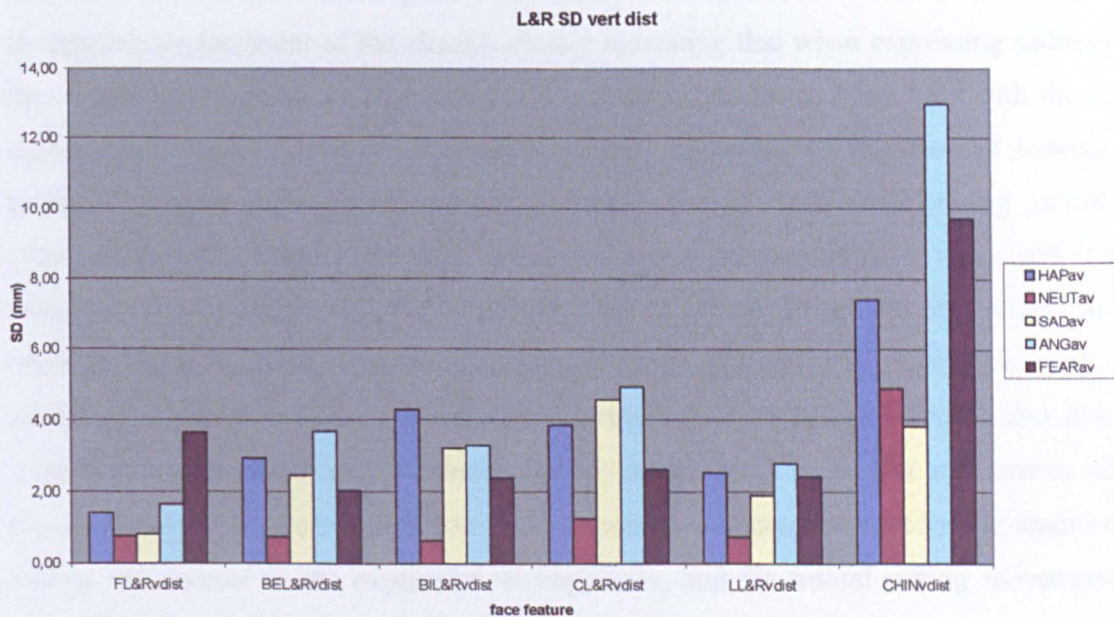
Figure IX - Left & Right side averaged vertical distance by the five conditions: Happiness, Neutral, Sadness, Anger and Fear

DISCUSSION

The values in the tables and in the graphs above provide important and sufficient empirical evidence to reveal clear differences between the five different conditions investigated. From all the conditions, the neutral one is the only one which presents almost no facial movement, except for a minimal gesture of the cheeks (probably resulting from a technical adjustment of the singing attitude) and for some major chin movements clearly connected with the articulation of the words of the song text – 'Mein Vater, mein Vater'. The movement of the chin is of course for the same reason the area of most movement, increasing as it goes from neutral to happiness to fear to anger. The movement of the chin decreases as it goes from the neutral to sadness condition. This result corroborates one of Bassilli's (1979) findings, i.e., that in the facial expression of sadness there is a slight upward displacement of the chin. The movements in the facial expression of sadness are connected with an inward and upward movement of the inner part of the eyebrows and with an upward displacement of the cheeks, clearly indicating that when expressing sadness the face contracts, reducing its overall surface. A parallel could be draw here with the vocal expression of sadness, that is once when singing with a vocal expression of sadness the laryngeal ventricle enlarges reducing thus the overall surface of the vibrating part of the vocal cords and therefore allowing the (expressive) production of a more soft, weak, breathy tone. Fonagy's, (1962) empirical investigation on glottal behaviour during emotional speech, seems to corroborate this suggestion. He found that by soft, weak, and whispered voice phonations the laryngeal ventricle was wide and enlarged, and that the vocal folds were thick and swollen in angry complaint. The widest movements of all chosen facial points are connected with the expression of anger, except for the inner brows raising movements in the expression of happiness, and the frontal raising movements in fear. The other widest movement in fear is connected with an upward displacement of the cheeks, probably generated by the contraction of the Levator Palpebrae Superioris, which raises upper eyelid, and sphincters of iris (see Figure X). In happiness the eyebrows are lifted, and the upward displacement of the cheeks and the lips is even greater than in fear. This can be understood by the action of the contraction of the Orbicularis Oculi raising lower eyelids and even by the action of zygomaticus muscles, which draw the angle of mouth backward and upward.
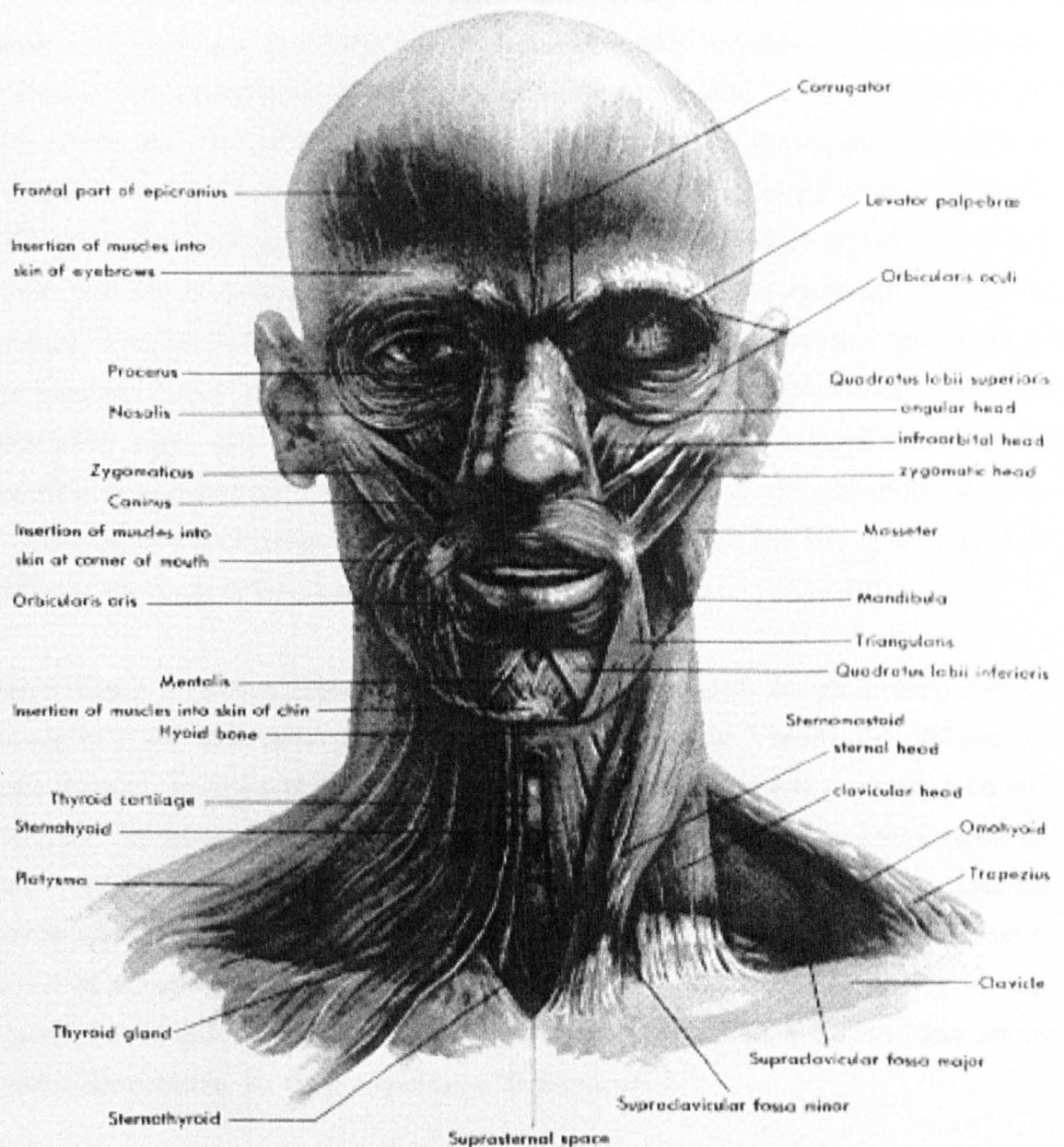
Figure X - Muscles of head and Neck, Front View. *Atlas of Facial Expression.* New York: Oxford University Press. Adapted from Peck (1987:7).

In happiness the cheeks' displacement is smaller than in the other two emotional expressions (anger and sadness) and wider than in the neutral condition. In sadness the movement of the lips is slightly narrower than in the other three facial expressions (anger, happy, and fear), revealing a downward displacement of the lips drawing in, and the corners of the mouth being depressed, by the action of the muscle Triangularis. Fonagy's

(1976) investigation of emotional articulation seems also to corroborate these findings. He found that anger was associated with violent movements between extreme articulatory positions, while disappointment was characterised by a relaxation of the tongue and of the soft palate, and decreasing speed of articulatory movement. According to Ekman & Friesen's (1978) manual for the Facial Action Coding (FACS), facial expressions imply different actions that can be studied and analysed. From the results of Experiment-I, it seems possible to conclude that singing with different emotional meanings involves the production of facial expressions, which imply a set of very different and specific facial movements. These movements are directly connected with the intended emotional expression sung, and can be detected and captured with the help of the suggested quantitative-objective procedure. This procedure seems to certify not only of the existence of these movements but also of the specific differences of their behaviour (according to the different intentions of the singer) independently of the perceiver's subjectivism.

Nevertheless, a question remained unanswered by Experiment-I: though it seems possible to conclude that these facial movements exist and that they are quantitatively different for each different condition of the sung phrase, is it possible to conclude anything about their meaning? Do these movements mean anything to someone? Do these facial movements and their particular differences have specific meanings that could be, for instance, communicated to an audience? And, if they do mean something, what sort of meaning could be communicated, inferred or recognised through the expressive display of these specific facial movements? These questions required an answer and, thus, they are the reason and necessity for the proceedings of Experiment-II.

4.3.2 EXPERIMENT-II: *MULTIPLE-CHOICE TESTING OF DESCRIPTIVE TERMS FOR THE PERCEPTION OF EMOTION FACIAL EXPRESSION IN SINGING, USING THE DYNAMIC INFORMATION OF POINTS OF LIGHT.*

According to Davidson & Correia (2002:242), it is nowadays well documented that audiences can detect "finely grained information about musical expression (timing, pitch, and dynamics modifications to structural features of the music) and intention (the emotional mood of the performer and the piece) from a musician's body movements when he or she is playing."

Aim

Based on this assumption and on Bassilli's (1978) work on the perception of emotion in the human face, and Davidson's (1991, 1993, 1994) work on expressive performance gestures a second experiment (Experiment-II) was carried out. The aim was to investigate if the abstract movement of the reflective markers on the singer's face filmed during the expression of the different emotional meanings had enough dynamic information available for an audience to detect, perceive and recognise the inner intention of each facial expression displayed (see Figure XI, for an example of a framed moment of the point-light movements display showing the emotional condition happiness). Linking these results to the results of the Experiment-I, it seemed that it would be possible to respond to the questions left unanswered at the end of Experiment-I.

Method

In this study the 15 observers worked a computer screen to view the display of the dynamic information contained in the movement of the markers filmed during Toni trials to express emotional meaning while performing Schubert's Erlkönig phrase: "Mein Vater, mein Vater".

Figure XI - An example of a framed image from the computer tracked facial movements, here showing 'Happiness'

For this experiment only one from each group of the five previous selected attempts for each emotion (see Experiment-I) was chosen for observation. This was the version witch the singer believed to be most representative of his/her intention. Thus, one version of each of the 5 emotional conditions (fear, anger, sadness, happiness and neutral) was displayed on the computer screen and the 15 perceivers had to rate among a list of seven different emotion descriptive terms (anger, fear, surprise, sadness, disgust, happiness, and neutral or no-emotion) which emotional expression was being displayed (see Appendix A). According to Robson (1993) who stresses the importance of minimising random 'noise' or bias in the data, more words than images were used to minimise the effect of a chance "guess" causing the result. Also, and to avoid the chance "guess", each emotional condition was displayed five times, in a random order.

Results

As can be seen in the Figure XII, the analysis of the raw data showed that from the display of the five emotional conditions in the point-light displays the perception of anger had the highest percentage of recognition (~85%), followed by happiness (~79%) and sadness (~72%). Fear (~59%) was the hardest of the emotional expressions to be recognised, being mostly confused with sadness, surprise, or even neutral. Nevertheless, the accuracy of recognition was way beyond the 50%. The neutral condition (~52%) was clearly harder to recognise because, in fact, it is a non-emotional state and it was mostly detected as fear, sadness or disgust. As shown in Figure XII, of the 15 participants that have observed the dynamic information displayed in the point-lights, 13 have been able to identify the expression of anger, 12 the expression of happiness, 11 of sadness, 9 of fear, and 8 were able to identify the expression of 'no-emotion'.
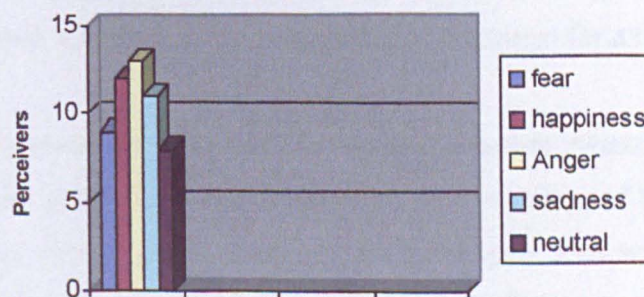

Figure XII - Accuracy of emotion recognition by point-light display

DISCUSSION

The analysis of the raw data indicated that the display of the point-light information in singing was sufficient to enable accurate perception. Also, it indicates that there is an equivalence of meaning between what is intended and what is detected. This indicates that also in the artistic domain dynamical information (movement) is an important perceptual source for emotion. This experiment (Experiment-II) might be seen as an important complement to Bassilli's (1978) investigation, because despite the relative perceptual subtlety involved in picking up information contained in his studies of the perception of emotion in the human face in everyday world, there has been a lack of information examining the artistic domain in which there are intentions of an expressive nature, which have to be contained within the constrained movements of technical and physical adjustments – as in the case of musical performance. It has been now possible to link the quantitative results of the Experiment-I to a reliable emotional interpretation of the point-light displayed information. So, it seems now possible to conclude that emotional facial expressions in music performance (at least the four basic emotions here under study) do not only have clear distinct movements in themselves and between one another, as the information they convey, even in the subtlest dynamic form, can be accurately perceived and recognised as categorical emotional meaning. However, despite the information contained in the point-light display being only visual, the study shows that the dynamic information is an important part of the communication's process in music performance (singing). Of course, singing as any musical performance is also made of acoustical elements (sound). These acoustical elements, despite their recognised significance, have not been yet investigated. So, further work is necessary to investigate the full process emotional communication in singing. Also, the way these two elements, visual and acoustical, interact in the communication's process of singing should be the target for a further investigation.

Bearing the limitations of the results above in mind, a further investigation was undertaken to provide empirical evidence on the perception and recognition of the different emotional meanings conveyed vocally and facially by both the singers. This was the proposed aim of Experiment-III: to investigate empirically and objectively how a certain audience perceives and recognises the emotions communicated through the voice and the face of the singers within the performance of singing.

4.3.3 EXPERIMENT-III: *MULTIPLE-CHOICE OF DESCRIPTIVE TERMS IN THE PERCEPTION AND RECOGNITION OF EMOTIONAL EXPRESSION IN SINGING*

Aim

Experiment-III aimed to study the perception and recognition of emotional meaning within the performance of singing, through the observation of videotaped performances of singers singing the same musical phrase with different emotional contents. So, the aim of this study was to investigate empirically and objectively how a certain audience perceives and recognises the emotions communicated through the voice and the face of the singers within different videotaped performances of singing (the different emotion examples of the videotaped performances of the two singers – Silvia and Toni - can be seen in CD 2).

Method

As in the first two experiments (Experiment-I and II), the selected musical phrase was "Mein Vater, mein Vater" from the Lied "Elrkönig" by Schubert. There were 5 stimuli for this experiment (one performance for each Emotion) done by the two singers (Silvia and Toni). For the recording, it was necessary to videotape the two singers in full-face close up. A professional Sony DSR 300 P-DVCAM video-camera and a professional post-production video table Pinnacle DC/DVD 2000 were used for the filming and video-taping, and the Betacam sound channel was used to input a spectogram through a special designed software – SpectraLAB–FFT Spectral Analysis System (16384 Size).

The audience task was to perceive and recognise the emotions, verbally and non-verbally, according to three different modalities: A. Acoustically (only voice), B. Visually (only face), and C. Visually and Acoustically (facial and vocal expression). In all the three modalities, the audience (15 participant observers) was asked to rate or label, verbally and non-verbally, the emotion perceived. In fact, the audience was asked to recognise, assess and classify which of the different emotions the performers were intending to express, and which among those expressions were perceptually more authentic (appropriate expression) or inauthentic (insufficient through to exaggerated expression) according to the intended emotional meaning perceived. According to the experimental design, the 15 observers were asked to rate among multiple descriptive terms which emotional condition (fear, anger,

sadness, happiness, and neutral) was being performed by both singers, and also to assess each performance according to the following bi-polar interpretation criterion: insufficient ← **to** → exaggerated (within a scale from 1 to 7, being '1' the mark to classify insufficiency and '7' to classify exaggeration). These data were obtained in two different ways, verbal and non-verbal (see Appendix B and C for an example of the computer screen lay-out, in Verbal mode and Non Verbal mode; see also the software especialy designed for these experiment in CD-4):

VERBAL – the perceivers should be able to rate the emotion perceived among a set of descriptive words-concepts (see Figure XIII), by picking up the 'term' that they considered to fit the emotional content of the facial expression displayed. These descriptive word-concepts were presented in a random order.

FEAR-ANGER-SADNESS-HAPPINESS-NEUTRAL

Figure XIII - Word-concepts for emotional rating

NON-VERBAL – the perceivers should be able to rate the perceived emotion among a set of still photographs taken from Ekman & Friesen (1976) universally recognised and tested emotion facial expressed examples (see Figure XIV), by picking up Ekman's facial expression that they considered to fit the emotional content of the singer's facial expression portrayed. These photos were presented in a random order.

Figure XIV: Facial expressions of five of Ekman & Friesen's (1976) basic emotions of still photographs' series. (Computer manipulated emotion in facial expressions from Calder et al., 1997, adapted from Bruce & Young, 1998:201).

In this context, Verbal means using word-concepts as those written in Figure XIII, and Non Verbal means showing pictures of the five emotional expressions using Ekman's (1976) still photographs of universally recognised facial expressions of basic emotions, and those can be seen in Figure XIV. The answers to the rating scales (Verbal and Non Verbal) were then used as sources of data for analysis.

Validation of the experimental stimuli

Both singers believed that their emotions were 'authentic' in that they were constructed out of their memories of those states. However, they were able to recognise that in their different interpretations some were more 'successfully' performed than others. Thus, from the fourteen different attempts at each emotional state performed, each singer chose his/her personally most satisfactory performance for each intended emotional expression, the one he/she considered to be the most 'authentic' one.

Before starting the analysis of the artistically performed emotions one last question should be answered, that of their empirical validation: Were these chosen examples of emotional expressions really corresponding to the expressions of everyday expressed emotions? In other words, were the chosen expressive examples 'authentic' enough? How could one be sure that the emotional content of those chosen expressive examples was really conforming to the emotional content of the same emotions in the 'real' world of everyday emotions?

In order to corroborate the singers' choice and to validate the emotional content of the singers' chosen performed emotions, two different ways have been followed:

- First (I), by comparing the acoustical analysis of these recorded emotional expression performed examples with the acoustical analysis of similar examples of expressed emotions in other existent studies on vocal emotional communication in speech, as well as in singing (Kotlyar & Morozov, 1976; Scherer & Siegwart, 1995; Sundberg, 1987; Scherer, 1995; Banse & Scherer, 1996);

- Second (II), by comparing the facial expressions of the two singers chosen examples of the performed emotions with the tested examples of the same emotions taken from Ekman & Friesen's (1976) still photographs series of basic emotions, and Izard's (1971) facial emotions photographs and charts (describing the facial features, movements, and muscles involved in each of the basic emotions), and Hjortsjö's (1969) work concerning the description of the visible appearance changes for each muscle of the facial expressive movements.

I - Validation through the visual inspection of spectral data

For the identification of the most relevant spectral parameters of the vocally performed emotion the SpectraLAB – FFT (16384) software was used:

> Software: SpectraLAB – FFT Spectral Analysis System
> FFT size = 16384
> Frequency resolution = 2,7 Hz
> Sampling Rate = 44,1 kHz
> Smoothing Window: Hamming

The three bars of the Schubert song were recorded using the Betacam sound channel and were then put through the computer software to produce spectral analysis of the different emotional interpretations. Figures XV and XVI show the different time spectogram for the different referred emotions for both singers: Silvia and Toni. Please note that spectograms are of course able to show many aspects of the vocal profile: onset, offset, fe/male vocal characteristics, and so on. In this validation, the focus is on the amplitude and harmonic content. The resulting spectogram profiles allowed the following sound analysis:

Emotion – Neutral: In this condition, the harmonic content of the two singers is very low, thus the singer's formant is not dominant, and the voice is weak in both amplitude and harmonics. The range of variability of F_0 (fundamental frequency) is small.

Emotion – Sadness: Here there is a relatively low harmonic content in both singers to that of the other emotions. With Toni's voice being far weaker in both amplitude and harmonics than in his other examples (around F_0). In both cases, the singer's formant is not particularly dominant (the strong harmonics are between about 2500 and 3500 Hz).

Emotion – Fear: Both singers show low harmonic content, with only the first two harmonics coming out strongly in Toni's case. Silvia is almost totally lacking the singer's formant, and her voice is very weak, the tendency to half voice/ whisper, seems to be how the impression of fear is given. For Toni, however, there is quite stronger harmonic content and voice amplitude than for Silvia, indicating a much fuller voiced interpretation (see explanation under in 'discussion of the 'eye-ball' validation').

Emotion – Happiness: Both singers show a much stronger harmonic content and amplitude, and the 2500-3500Hz area is stronger than in the fearful example. The distribution of the energy in the frequency spectrum vocal energy is concentrated around the Fundamental frequency (Fo) and the second harmonic. The singer's formant in both singers is not as strong as in Anger's example, but is stronger than in all other cases.

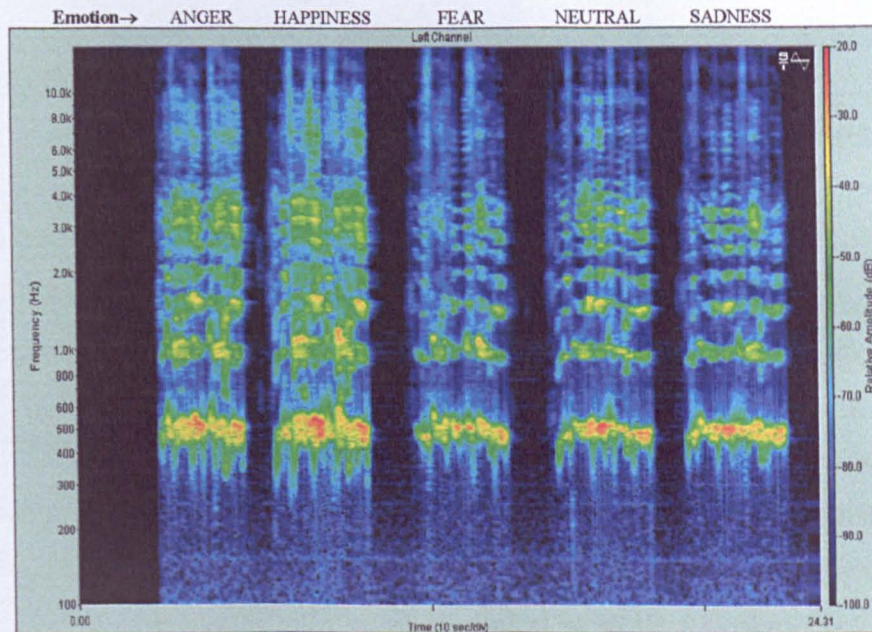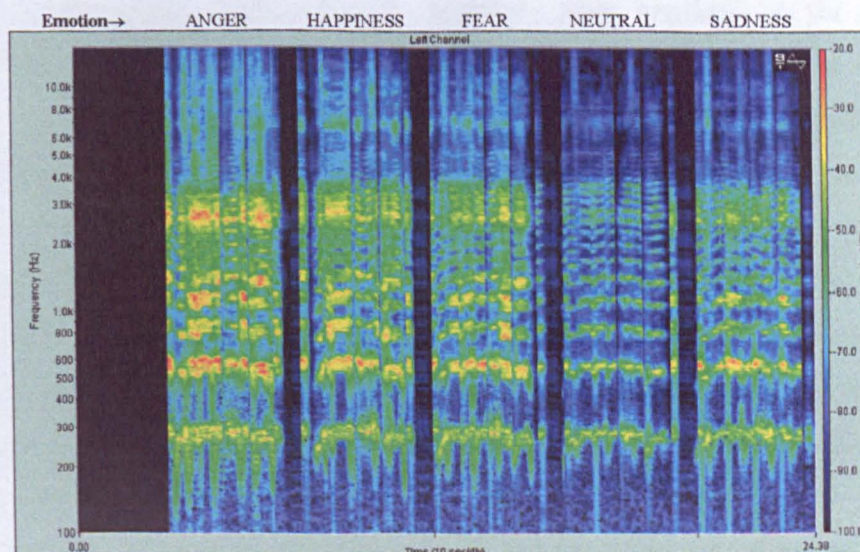**Figure XV** - Time spectogram of Silvia's voice

Figure XVI - Time spectogram of Toni's voice

Emotion - Anger: This example showed the most dramatic change in both amplitude and harmonic content, which are considerably greater for Silvia and Toni. There is a greater amount of energy in the frequency spectrum around Fo and the next three harmonics for both singers. The singer's formant, in case of Toni rounds the 2.900-3.000 Hz, and in case of Silvia is higher, around 3.400 Hz. The overall amplitude of the voice in both singers is stronger than in all the other emotions.

Discussion of the acoustical validation

The analysis of the acoustical voice parameters seems coincident with analyses already existing on vocal emotional expression in singing (see, for instance, Kotlyar & Morozov, 1976 and Scherer & Siegwart, 1995). This suggests that the performed emotions of the different examples match the intended emotional content within the singing performance context. These results are also similar to the empirical evidence found in the study of the acoustical parameters of voice when the referred emotions are expressed in speech (see Lieberman & Michaels, 1962; Sedlácek & Sychra, 1963; Williams & Stevens, 1972; Fonagy & Magdics, 1963; Sundberg, 1987; Scherer, 1991, 1995; and Banse & Scherer, 1996). This suggests that the performed emotions of the different examples match the intended emotional content also in everyday world contexts. Of course the fact that the singers were singing different pitches (Silvia's voice pitch was around A3/B flat 3 and Toni around an octave lower, around B2/C3) certainly have implications for the different amplitude and harmonic contents. For instance, in case of the female singer, because Fo is almost an octave higher, for the same frequency region the density of partials (i.e., the number of partials for the same pitch interval) is smaller originating less definition of the harmonic content and so this explains why the harmonic content on Silvia's voice spectrogram reveals a much less clear harmonic content profile in the different emotional versions. Also, it should be added that there are some main differences between male and female in the singer's formant. Sundberg (1987:101) explains, for example that the fourth formant affects the personal component of the sound of the voice because the main factors contributing to the frequency of this formant are "the vocal tract length and the tract dimensions within and around the larynx tube", and the relevance of one or both of these factors depends "on the cross-sectional area of the lower part of the pharynx" (Sundberg, 1987). Of course, and according to considerations such as these pointed out by Sundberg,

some differences are to be expected between the two singers even if both are singing with the same emotional content. On the other hand, Bloothooft (1985) has shown that the pitch per se does not have a clear effect on the level of the singer's formant (the only that could be ascribed would be the fact that very high notes are usually sung louder). Also, according to Sundberg (1987), in soft phonation, the fundamental is the strongest partial in the spectrum (confirmed here by the expression of the neutral emotional condition); in a moderate version, the strongest partial would be the closest to the first formant (what seems to be corroborating the expression of sadness); and as loudness increases, the higher the spectrum partials gain more amplitude than the lower spectrum partials (what can be observed as it goes to more full voiced emotions like fear, happiness, and anger).

II - Validation through the visual analysis

For validation of the emotional stimuli, i. e., the expressive facial gestures produced by both singers while performing the Schubert phrase with different emotional intentions, a carefully analysis process of the different performances was necessary. Since this investigation was relying on the performers interpretations and some validation of differences in their emotional contents, an "eye-ball" study of the muscle movements was undertaken, comparing the collected data of the singers' performances:

a) with those of Hjortsjö's (1969) study of the effects of contractions of facial muscles (see Figure VIII);

b) with those universally tested Ekman & Friesen's (1976) still photographs series of basic emotions (see Figure XIV and XVII), and with those of Ekman & Friesen's (1978) study of the effects of contractions of facial muscles in the manual for the Facial Action Coding (FACS);

c) with those of Izard's (1971) facial emotions photographs and charts (describing the facial features, movements, and muscles involved in each of the basic emotions);

d) and with those where a more intuitively artistic observation of the facial emotional expression was used – Peck's (1951) *Atlas of Human Anatomy for the Artist*, and Peck's (1987) *Atlas of Facial Expression*.

Figure XVII (below) compares the facial expression of the different emotions of the framed facial emotional expression peaks of the two singers (Toni and Silvia) with the universally tested Ekman & Friesen's (1976) still photographs series of basic emotions:

NEUTRAL	FEAR	ANGER	HAPPY	SAD
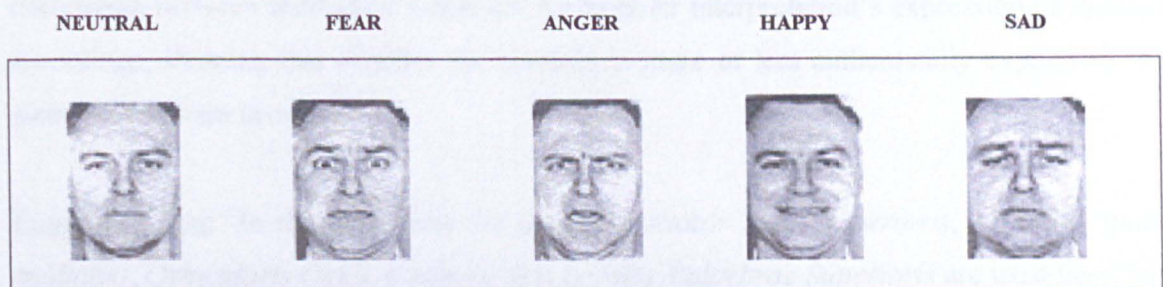
Ekman & Friesen's (1976) examples of emotion facial expressions

Toni's emotion facial expressions

Silvia's emotion facial expressions

Figure XVII - Comparing singer's facial emotional expressions with Ekman & Friesen's (1976) still photographs of universally recognised facial expressions of basic emotions.

From a brief examination of the Figure XVII it is apparent that Toni and Silvia were using the facial muscles appropriate to each emotion expressed. The results of the carefully undertaken 'eye-ball' analysis indicated that (see Figure X for a general view of the facial muscles):

Emotion – Neutral: In this condition, the movements show a very limited range of muscle activity, with a high degree of correlation between the two singers' use of their faces in this condition.

<u>Emotion – Sadness</u>: In this condition the muscles *Corrugator*, *Triangularis*, *Depressor Supercilii*, *Mentalis* and *Depressor Labii Inferioris* are used extensively, with Toni showing the most movement activity here, and Silvia a moderate range of activity. There is a correlation between individual's data for the bi-polar interpretation's expression of sadness recordings, showing that whether the emotion is more or less authentically expressed, the same muscles are involved.

<u>Emotion – Fear</u>: In this condition the muscles *Levator Labii Superioris*, *Frontalis (pars medialis)*, *Orbicularis Oculi*, *Caninus*, and *Levator Palpebrae Superioris* are used here, but the degree of involvement varies according to individual (and bi-polar) interpretation. For instance, in case of Toni both lots of muscles are equally involved in all interpretations of fear (see explanation under in 'discussion of the 'eye-ball' validation'). For Silvia, there is little *frontalis* activity, but more in her more authentic interpretation of fear.

<u>Emotion – Happy</u>: In this condition the involved muscles are: *Zygomaticus Major* and *Zygomaticus Minor*, *Risorius*, *Orbicularis Oculi* and *Frontalis*. Like in neutral, there is a high degree of correlation between singers' interpretations.

<u>Emotion – Anger</u>: In this condition the involved muscles are: *Platysma*, *Procerus*, *Depressor Labii Inferioris*, *Mentalis*, and *Levator Labii Superioris*, and *Aleque Nasi*. Here, both singers use very similar formations and degrees of activity in the several renditions of the emotion.

The aim of this, of course, was in itself a part of the empirical investigation undertaken in Experiment-III (see below results of Non Verbal statistical analysis of Experiment-III).

Discussion of the 'eye-ball' validation
The results of the 'eye-ball' analysis seem not only to confirm that an appropriate set of muscles was being used for each expressed emotion, but also showed that different groups of muscles and of muscle actions were active according to the expression of the different investigated emotional conditions. By having explored the visual stimuli through these comparative means, it seems possible to conclude that the same muscles or similar group of muscles actions are being used and can be recognised and identified in both sets of images:

the images of the singers' peaks of emotional facial expressions and Ekman & Friesen's (1976) still photographs series of basic emotions. It seemed thus that according to the positive results of the preliminary visual analysis undertaken, it would be useful to submit the videotaped sung phrases to further and more scientific visual analysis.

Preliminary conclusions through the validation of the performed stimuli

In summary, through the validation analyses undertaken, it has been possible to suggest that the acoustical content of the sung phrases is coincident with the acoustical content of the analysis of singing with emotional content (see Kotlyar & Morozov, 1976, and Scherer & Siegwart, 1995), and even with the analysis of speech with emotional content (see Lieberman & Michaels, 1962; Sedlácek & Sychra, 1963; Williams & Stevens, 1972; Fonagy, 1976; Sundberg, 1987; Scherer, 1991, 1995; and Banse & Scherer, 1996). Equally, it has been possible to conclude that the visual information conveyed in the singers' facial expressions is very much coincident with the visual information of the facial expressions of four of the basic emotions (and the neutral condition) taken from examples of Ekman & Friesen's (1976) still photographs series of basic emotions. Thus, it is possible to say that the emotional content of the performed phrases by both singers seems to belong to the same meaning domain as the examples used to validate them. This is, of course, what Experiment-III was supposed to investigate and, therefore, to confirm. In fact, the main purpose of Experiment-III was, as said, to investigate if, and how, the facial movements and the vocal sounds of these 5 sung examples are recognised as expressive of the 5 intended emotional conditions.

These preliminary analyses (acoustical and visual) to Experiment-III reveal, also, that both larynx and face move and work very differently according to each emotional state. The particular profiles of each share some common characteristics. For instance, when singing with a sad expression, the face contracts, reducing its overall surface area, and so too does the vocal sound, producing a more breathy, whispered tone, indicating that the ventricle folds are further apart (and thus the vibrating surface of the vocal folds is also reduced) than in the other emotional conditions. A similar result was obtained for the communication of fear: whereas Toni used a much fuller sound and the *Frontalis* was high lifted as is

typically used in frightened facial expressions in everyday world (Ekman & Friesen, 1976 and Izard, 1971), Silvia used a much lesser activity in this muscle. This suggests that the two singers had slightly different interpretations of what kind of fear was being performed and how it was created. One possible reason for this difference might be that Toni's interpretation introduced the emotional element of terror, (which, within the 'Erlkönig' song itself, is surely a part of the emotion of the child when he, suddenly and in a fearful state, calls out for his father by noticing the terrifying presence of the ghost of the Erlking). Silvia, otherwise, said that she had focused more on a generic expression of fear, and did not have the element of terror in mind when singing. Thus, this may have been why in her performance the muscle *Frontalis* was not so active, and the voice had lower sound level than Toni's vocal interpretation.

Whilst the raw data show that there was a correlation for all emotional states, especially for happiness, it is worth mentioning that the two singers noticed discomfort in their throats when singing in the angry condition. Upon this preliminary analysis of the face, it seems that this effect could be attributed to a harder functioning of the muscle *Platysma*, which is definitely involved in the expression of anger, but which may function as an intrusion for the required adjustments for the singing technique, once the use of this muscle implicates the tension of the tongue and jaw muscles preventing in this way the larynx from working and vibrating with the singing required degree of freedom. So, something a singer should definitely be aware of when trying to express anger within the performance of singing.

To conclude, the results here just presented add to Experiment-III by showing that the emotional content of these sung phrases does match the facial features and the acoustical profiles of other authors' investigations on the same subject. Of course, without the statistical differences in these more or less successful interpretations, it is only possible to begin to theorise about what might allow for the differentiation between the interpretations to occur. It appears that there are qualitative differences in the intensity of how the muscles are used. That is to say that if the singer is clearly working with a stronger inner intention the effect is more successful.

THE STATISTICAL DIFFERENCES (Experiment-III)

Analysis of data resulting from the observers' choice among multiple emotion terms to describe the performance

The data from Experiment-III, as a result from the answers of the 15 perceivers to the singers' performances stimuli according to the referred rating scales (verbal and non-verbal), were used, as said, as sources of data for analysis. These data were analysed using two separate repeatedly measured analyses of variance in order to explore differences in mean scores across conditions and between respondents, who had viewed and rated the performances. The emotions presented were (Sad, Happy, Angry, Fearful, and Neutral) and respondents were asked to make their judgements according to different Stimulus Presentations (Vision only, Sound and Vision, Sound only) and two different Modes of Response to the stimuli (the Verbal response and the Non Verbal response). Additionally, as referred to above, there were two different performers on whose performances the respondents had to base their judgements.

Analysis of correct identification of the emotions sung

The first analysis undertaken was to explore how 'accurate' respondents were in recognising the emotion being portrayed according to the type of stimulus presentation. The number of correct responses for each experimental condition was 5 (the four basic emotions, plus a neutral emotion option). These were then analysed using a 3-way analysis of variance: 3 (Stimulus Presentations) x 2 (Response Modes) x 2 (Performers). The analysis revealed that the type of Stimulus Presentation and the Stimuli Mode both appeared to affect the accuracy of the respondents' judgement of the emotions, although the effect of Mode of Response was not quite significant [Main effect of Stimulus Presentation $F(2,28)=12.66$, $p<0.001$; Main effect of Mode of Response $F(1,14) = 3.973$, $p = 0.06$]. Tables III & IV show the mean rating of correct responses according to Stimulus and Response type. It appears that the easiest stimulus to respond correctly to is a sound and vision stimulus, followed by a vision only, and then a sound only stimulus (which seems to corroborate Ohgushi & Hattori's (1996a), and Salgado (2000) empirical findings on emotional communication in performance of vocal music). It also appears that more accurate responses are given when the Mode of Responses is Non Verbal rather than Verbal.

Table III: Mean rating and standard deviations of correct responses by Stimulus Presentation

Presentation	Mean(1 – 7 expression)	sd
Sound & vision	4.50	0.10
Sound only	3.60	0.17
Vision only	4.20	0.11

Table IV: Mean rating and standard deviations of correct responses by Mode of Response

Mode	Mean(1 – 7 expression)	sd
Verbal	3.90	0.09
Non Verbal	4.20	0.11

Analysis of response times

In order to investigate whether the processing of the stimulus information varied in duration, a 4-way repeated measures analysis of variance was carried out on the response times. There was no overall difference in the response times for Verbal versus Non Verbal Modes of Response [$F(1,14) = 1.32$, ns], nor was there a difference depending on the individual Singer [$F(1,14) = 0.85$, ns]. There were, however, overall differences in the response times according to the emotion being conveyed [$F(14,56)=3.331$, $p<0.05$], with participants responding more quickly when the emotion conveyed was Anger and more slowly for Fear (see Table V). Differences in response time were also found depending on the Stimulus Presentation [$F(2,28)=27.664$, $p<0.001$] with participants taking longer to respond to 'sound only' stimuli (see Table VI). There were no significant interactions between these effects.

Table V: Main effect of Mean response time according to Emotional condition

Emotion	Mean (in secs)	sd
Happy	3.30	0.38
Sad	3.70	0.64
Angry	2.80	0.22
Fearful	4.90	0.70
Neutral	3.50	0.55

Table VI: Mean response time for each type of Stimulus Presentation

Stimulus type	Mean (in secs)	sd
Sound & vision	2.40	0.15
Sound only	5.80	0.71
Vision only	2.80	0.27

Analysis of the observers' ratings of goodness of fit of the Emotion expressed

There was little variance in the scores away from the middle of the rating scale suggesting this was not a sensitive measure. However, a 4-way repeated measures analysis of variance

was still carried out on the ratings (1-7). There was no overall difference in the response times for Verbal versus Non Verbal Modes of Response [F(1,14) = 0.21, ns]. However, there was a difference depending on the Singer [F(1,14) = 5.85, p<0.05] with Silvia having a slightly mean lower rating (3.6) than Toni (3.8). There were overall differences in the ratings according to the emotion being conveyed [F(14,56)=10.74, p<0.001], with participants giving a higher rating when the emotion conveyed was anger and lower for fearful (see Table VII). Differences in response were found for the Stimulus Presentation but these were not quite significant [F(2,28)=2.77, p= 0.08], with participants rating 'sound only' slightly lower on the scale (see Table VIII). In addition to the above main effects, there was also a significant interaction between response and singer (see Figure XVIII). While Silvia got higher ratings for Verbal responses than Non Verbal responses, for Toni it was the other way round.

Table VII: Mean ratings for Emotions averaged across conditions

Emotion	Mean (expression rating)	Sd
Happy	3.70	0.11
Sad	3.70	0.08
Angry	4.10	0.12
Fearful	3.30	0.11
Neutral	3.50	0.13

Table VIII: Mean rating for each type of Stimulus Presentation

Stimulus type	Mean (expression rating)	Sd
Sound & vision	3.80	0.12
Sound only	3.50	0.08
Vision only	3.70	0.09
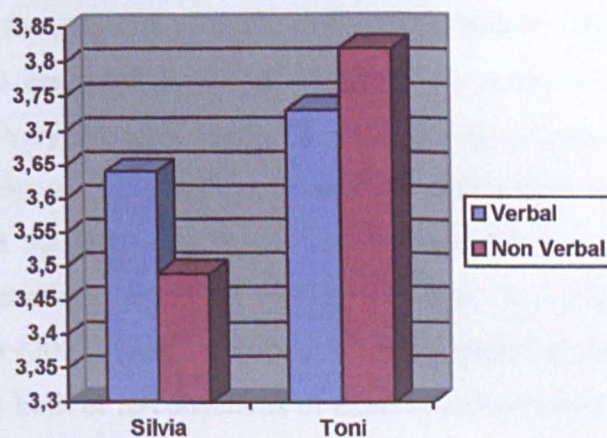
Figure XVIII: Interaction of rating according to singer and type of response

Discussion of the findings

According to the empirical evidence and the results of the statistical analysis in Experiment-III it seems possible to conclude that sound and vision together provide the most accurate stimulus presentation for the judgement of emotion in the two singers (Table VI). Visual information alone is a more accurate means in the expressive communication of emotional content in the performance of singing than acoustical information alone. Not only did visual information get higher ratings than acoustical information, visual information mode was more quickly responded to by the judges (Table IV). These findings corroborate the results of Ohgushi & Hattori (1996a), and Salgado (2000). However, the stimulus presentation providing the quickest answer from the audience was sound and vision together.

Another conclusion to be drawn is that, there is a strong and parallel link between visual and acoustical communication of emotion. Also, from the validation process in Experiment-III, it could be concluded that even the least expressive attempt to convey emotional meaning showed, though with less activity and accuracy, related acoustical and visual profiles. This finding might be interpreted, according to singer's reports on emotional expression in singing (at least for the basic emotions studied here), as the clear indication of emotional content while singing. Concurring with these conclusions, Runeson & Frykholm's (1983) research on expressive behaviour explored the nature of physical gestures and presented empirical evidence showing that covert mental dispositions are specified in movement and can be detected by observers. A relevant conclusion to be drawn from Experiment-III is that, vocal and facial expressions used by singers when trying to communicate emotional meaning, though allowing for the necessary vocal and musical adjustments, are in fact very similar to the vocal and facial expressions used for emotional expression in the everyday world. The singers' reports confirm this and the findings of this experiment suggest that the highest ratings of the audience answers be connected to the authenticity of the emotions portrayed by the singers. A subsequent conclusion, corroborating Persson's (2001) research, is that the 'performed' emotion is not a 'faked' emotion but a construct built of re-collections of experienced emotions. Non Verbal Mode of response is, when considered globally, a more accurate way of labelling the expressive communication of emotional content in the performance of singing. Rather intriguingly,

however, Silvia's performances were rated higher in Verbal Mode, and Toni's performances in Non Verbal Mode. This might, of course, be connected with personal and idiosyncratic characteristics of the singing of each performer, but the interaction of rating according to singer and type of response shown Figure XVI suggests that Toni performed each emotion 'better'. The mean ratings for emotions averaged across conditions (Table V) showed that <u>fearfulness</u> was the most difficult emotion to recognise in singing, followed by the <u>neutral</u> condition. Immediately after, <u>happiness</u> and <u>sadness</u> were equally well recognised, and the easiest emotion to recognise was <u>anger</u>. The main effect of mean response time according to emotion (Table III) showed that <u>anger</u> was the most quickly identified in singing. <u>happiness</u> was the second emotional condition to be recognised, followed by the <u>neutral</u> condition, and then <u>sadness</u>. Finally, the least quickly identified emotional state condition to be recognised was <u>fearfulness</u>.

Experiment-III has investigated how accurately facial and vocal emotional expression is labelled and assessed in the performance of singing. Emotional expression in singing has been labelled and assessed according to different modes (Verbal and/or Non Verbal), according to different stimulus presentation (Sound and/or Vision), and to different differential parameters (1, not sure of the emotion ← to → 7, definitely the emotion). The relevance of the results and findings in this experiment definitely seems to concur with some major conclusions concerning how an audience perceives and recognises the emotional meaning expressed in the performance of singing. Experiment-III was however processed and conducted in a laboratory. So, the aim of the Experiment-IV was to bring the research from the laboratory into the 'field'. How reliable are the findings of Experiment-III when submitted to an ecologically valid investigation? How accurate is the performer-audience emotional communication in the performance of singing under realistic conditions? Is the emotional meaning accurately perceived and recognised in a real life performance as it is in the laboratory conditions? Is the emotional meaning of the musical structure consistently vocally and facially communicated when performing a full song under the realistic conditions of a concert room? The practical aim of Experiment-IV, thus, was to investigate how to elaborate an expressive tool, based on a model of a real-time life-performance feedback, that might help singers, singing teachers, and singing students to test and improve the expressiveness of their performances.

4.3.4 **EXPERIMENT-IV:** *A REAL-TIME LIFE-PERFORMANCE COGNITIVE FEEDBACK STUDY*

Aim

The hypothesis underpinning Experiment-IV was: if, an understanding of the degree of accuracy of the perceivers' recognition of the singer's emotional interpretation could enable the performer to have a clearer perception of his/her emotional expression and, consequently, to help him/her to improve the expressiveness of his/her performance. If this hypothesis was successfully tested, it might be possible to conclude that the process used in this experiment could aid other singers and singing students to improve the way they use their voices and their faces when interpreting emotionally the structural content of a musical piece.

In Experiment-III, as it has been above reported, some attempts at emotional expression are considered more 'authentic', i.e., more successfully performed, than some others. Thus, it seems important that in the performance of music in which structure clearly points to an interpretation of categorical emotional meaning, the singer might be able to convey that meaning accurately by using facial and vocal cues to contribute positively to the communication of the intended emotional content. So, it seemed relevant to elaborate a 'tool' that could give immediate cognitive feedback of the degree of expressiveness of a real-time performance, i.e., an "expressive tool" where the singer's ability to express emotional meaning could be tested and improved. This has been simultaneously the aim and the process of the Experiment-IV: the elaboration and testing of a reliable cognitive feedback for a real-time life-performance that could function as an "expressive tool" for the use of the singers, singing teachers and singing students, which might help them to test and to improve the expressiveness of their performances.

Before the main experimental procedure was undertaken four steps were considered:

1. The emotional meaning of the song was explored through an analysis of the musical structure of the song (see Appendix F), alongside a consideration of the meaning of the Lied's poem and the general poetic context of the song cycle.

2. Computer software was designed to create a reliable way to register the audience feedback to the performer's interpretation. See Appendix G ('Software

Programme Text' for Experiment IV) + CD 4 (software designed especially for the Experiment IV).

3. A performance was worked on so that the emotional content found in the musical-poetic analyses could be developed. See in CD 3 (the videotaped performance-trials).

4. An experimental context was developed, where the planned performance, like in a real time concert situation, was assessed through an audience's real-time feedback. Therefore, and according to Emery Schubert (2001:409), this experiment might be considered as an 'ecologically valid' experiment, that is the listeners were listening to an uncontrolled, real piece of music rather than an excerpt or controlled auditory stimulus.

The development of the experimental-study

As mentioned, the first issue to be considered was a software design to reliably register the audience's categorical recognition of the performed emotions. The methodology devised aimed to create a situation of an 'almost' continuous experiment. Of course, as Schubert (2001:395) explains, "categorical responses require the participant to select one or several answers from a range of answers. (...) In a continuous response task, this may either interfere with the listening experience, or lead to the omission of some of the responses or guessing so as to minimise interference." Nevertheless, the target of this experiment was not to investigate all the unfurling emotional continuous nuances that the musical structure of the song might eventually convey. In fact, Experiment-IV aimed to investigate if the performer, according to the analysed structural meanings and within an ecologically valid framework, was able to communicate accurately, during the continuous development of the musical structure of the song, all the previously selected emotional contents (see Appendix F). Therefore, categorical responses were made, but over time, and according to intentional peaks of emotional expression performed, which had been programmed and rehearsed according to the analysis of the musical structure and the interpretation of the song.

In this experiment, the performer aimed to convey four different emotional peaks, which corresponded to the piece's musical structure and the four emotional contents (happiness, fear, sadness, and anger). In this experimental study, though the important features of the

musical structure have been carefully observed by the performer, what has been judged as 'emotional stimuli' was not the musical structure itself, but the singer's performance of the song, i.e., the singer's capacity to convey through the expressive vocal and facial cues, the four most relevant emotional contents analysed in the music structure and in the poem of the song. So, the main questions to be answered in connection to this experiment were:

How, or by which means, did the performer decided to convey the chosen emotional meaning? Which, musical elements and expressive cues were used by the performer to convey the emotional meaning? To what extent was the audience able to rate the performer's expressiveness? And, how (or if) the performer has been able through the audience's feedback to improve the expressiveness of his performance? Therefore, the aim of this experimental study was, as said, to validate and improve the performer's communication through the audience's identification and recognition of the emotional musical contents portrayed during the performance. The audience responses were made through the use of the software package described below and questionnaires used to provide a more detailed qualitative type of responses in order to obtain as detailed an understanding as possible of the experience for them.

Materials used

For the development of the experimental procedure following software and hardware used were:

Software:	Macromedia Director 8.0 Xtra - FILE I/O
HardWare:	Multimedia Personal Computer IMAC (for the development of the Macintosh version)

For the experimental procedure itself, the material used consisted of:

Infrastructure:	ETHERNET (10 Mbits) with limited access to the 4 computers involved in the experiment. Hub 8 doors
Server:	Multimedia Personal Computer (responsible for generating the logs in real time and for the laying out of graphs with the audience's answers)
3 distribution posts: (each one comprising)	Multimedia Personal Computer IBM 5 Samsung Monitors 5 Sony Stereo headphones

The audience questionnaire

This was developed along the lines described by Smith (1995) who argues for an Interpretative Phenomenological Analysis (I.P.A.) of meaning as well as the employment of more objective means, such as the physical measures used in earlier experiments. I.P.A. depends on qualitative data collection techniques so that the investigator is able to explore the respondents' own experience as closely as possible. This was attempted here by developing a questionnaire (see Appendix D which asks "open'ended" questions, thus enabling the responses to be as personalised as possible). Consistent with I.P.A. and other qualitative questionnaire work (see Robson, 1993, for example) a thematic analysis was then undertaken looking for emergent themes. The results are explored below, alongside a "talk-aloud" critical reflection which was made by the singer himself in order to explore the extent to which he was able to respond successfully to the empirical technique as well as the results obtained.

Method

Toni learned and sang the Lied 'Die Post' from Schubert's cycle 'Die Winterreise'. As mentioned, the intention of this investigation was also to serve as a pilot-study for the interpretation of the Schubert's cycle in his PhD recital. Therefore, the Lied was not only analysed structurally, but as an integrated part of the general context of the proposed song cycle (see Appendix F). The intended interpretation of the Lied was then communicated to the piano accompanist. The musical moments where the emotional peaks were 'supposed' to happen (that is, structural features indicating a specific emotional response – see Cook & Dibben, 2001) were linked to the places in the musical structure where the pianist underlay the correspondent features.

Despite the important role of the pianist in the communication of the musical structure, it was decided for the sake of the experiment that the pianist should try to stick to the singer's interpretation and to avoid any other role that may influence the responses of the audience.

She aimed therefore to underlay the selected musical features at the previously chosen moments, holding back her playing, and leaving the main role of the interpretation of the song to the singer.

Fifteen participants were asked to observe performances of the song and to judge the musically conveyed emotional meanings. They were seated at computer screens, and they were asked to rate in real-time the emotional content, or the different emotional contents, occurring during the performed interpretation by pressing emotion buttons indicated on the computer keyboard. The choice of the descriptive terms was also limited to the four main emotions found in the musical-poetic structural-interpretative analysis of the song (shown in Appendix F). There were three groups of five screens. Each group of five screens, for technical reasons, was connected to a distribution post. This clustering point is represented by a colour. So, there are three colours shown in the two figures: red, blue, and green. So, that the answers of each group of five observers appeared in a graph lay out, as red, or blue, or green lines. Therefore, the first group had five lines coloured in red. The second group five in blue, and the third group five in green. When no colour appeared on the graph that means that no answer was given, thus no emotion button was pressed. The song was performed twice, and on each occasion the audience was asked to rate the emotional meaning over the course of the performance. Each listener pressed the key according to the emotion perceived previously chosen (S-sadness; H-happiness; A-anger; F-fear). The answers were given in a real-time during the performances, and the emotion button was held down as long as the emotion was perceived. After the interpretations, the audience members were interviewed individually to investigate their reactions to the study. Toni used the time after this to make his own critical "talk aloud" responses to the investigation.

Experimental Procedure

After the first performance, the total number of answers rated was summarised and a graph of the performance was printed out (see Figure XIX). This graph showed a profile of all the audience selected emotions, i.e., a kind of report of the audience's emotional meaning perception and recognition, and the corresponding timing (in seconds) of the exact place within the musical structure where the emotion was recognised. This print out was used by the performer as a means of rectifying the extent to which the interpretation he 'thought' he had given was communicated to and detected by the audience.
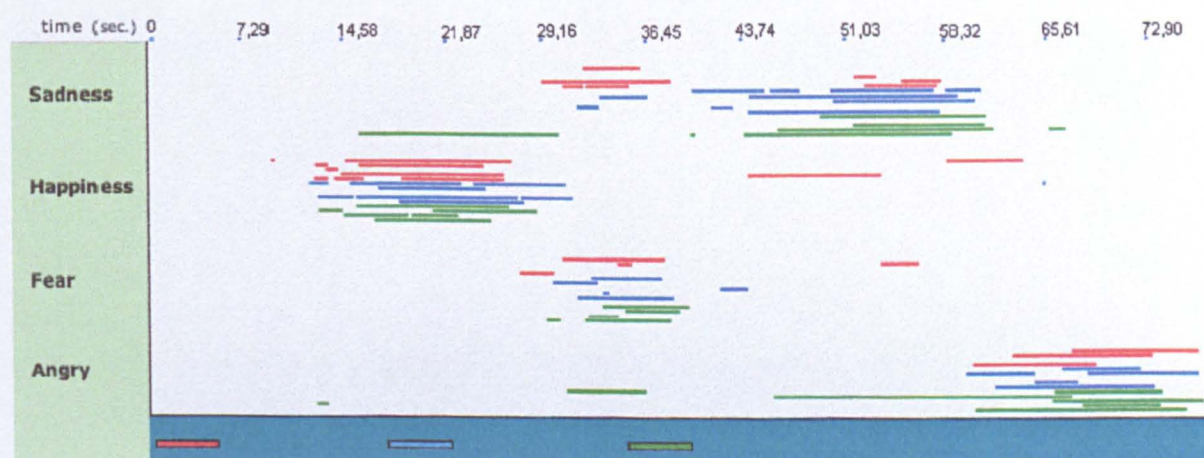
Figure XIX - Profiles of the perceived emotional meaning(s) of the 1st performance of the
Schubert's Lied 'Die Post'

Description of the performance event

Schubert's Lied 'Die Post' – a pilot study for the interpretation of Schubert's cycle 'Die Winterreise'

Pianist and singer started the performance with the piano introduction part being played on a full size Steinway, by a professional female pianist, aged 32 (Xao Li). At bar 9 Toni started to sing attempting to express vocally and facially the excitement and expectation of the 'character' of the song indicated in the arrival of a mailman bringing a letter with news from a belovèd. So, the vocal and facial expressiveness of expectation and light excitement of the beginning of the song developed to a new expressive and concrete aim, which appeared clearly and strongly with the first emotional peak at the beginning of bar 15. This emotional peak corresponded to Toni's attempt to express vocally and facially the emotion of happiness expressed in the words *mein Herz*, and with the 6th major jump (A flat 3 → F 4), being a structurally salient moment, here indicating happiness as in many other cases in

the music literature (see Cooke, 1959: 65). Also, since it is a high-pitched and ascending interval, it could in this context be indicative of happiness (see Maher, 1980 and Gerardi & Gerken, 1995). According to the performer's interpretation of Schubert's instructions at the beginning of the song *Etwas geschwind*, this moment occurred around the 20th second of the performance. Also, according to the experimental instructions given to all the participants in the audience at the beginning of Experiment IV, the 'emotion-button' should have been pressed as soon as an emotion was perceived and as long as the emotion remained perceptible. So, at this moment (around the 20th second) the audience should have given (in the best case) clear emotions of expectation and joy by pressing the happiness button, among the choice of the four descriptive terms. Different expressive attempts followed, corresponding to the emotional peaks previously and carefully planned and rehearsed according to the performer's structural and interpretative analysis of the song. These moments occurred around the 35th second, the 50th second, and the 70th second of the performance. The full description of the structural and performative analysis of the Lied *Die Post* used as a support to Toni's interpretation of the song in Experiment-IV can be found in Appendix F, where a further discussion of the most relevant elements of the relationship between the performance vocally and facially expressed emotions, the audience's perception of the emotional meaning, and the structural features of the music, is presented.

The audience's real-time recognition of the different performed emotional meanings was registered again and again as raw data for feedback on a computer-generated graph (Figure XIX). The singer, after the first performance, had access to this graph and carefully examined it. The singer could also use the real time videotaped performance to observe and analyse in playback his first performance (see CD 3 for an example of a videotaped performance-trial). This observation process took a total of 3 minutes. According to the interpretation of these first results, and by comparing them with the final conclusions of the performative analysis (presented in Appendix F), the singer could study and understand which expressive cues were missing from his performance of the song, facially as well as vocally, and those that could be improved or performed in a clearer manner in order to get from his audience a better understanding of the emotional meaning he intended to express. Therefore, the missing or confusing features of his vocal and facial interpretation were identified, corrected, and improved during the second version.

A second performance followed, then, where all the new, or the improved expressive cues, were added to the interpretation, and again the audience rated the intended emotional meaning, according to the same parameters of the first performance. The new data were printed out on a second graph (see Figure XX), and the singer was able to compare the new results with the older ones, to check if, in the second performance, he had communicated more accurately the different emotional contents of the performed song. It should be added that, though there were two different performances of the same song, the intended interpretation of the song was the same on both occasions. The only change was at the level of the performance's accuracy, that is if the singer was able to communicate more or less accurately the same intended emotional meanings (see Appendix F for information on the more accurate rendition of the song).
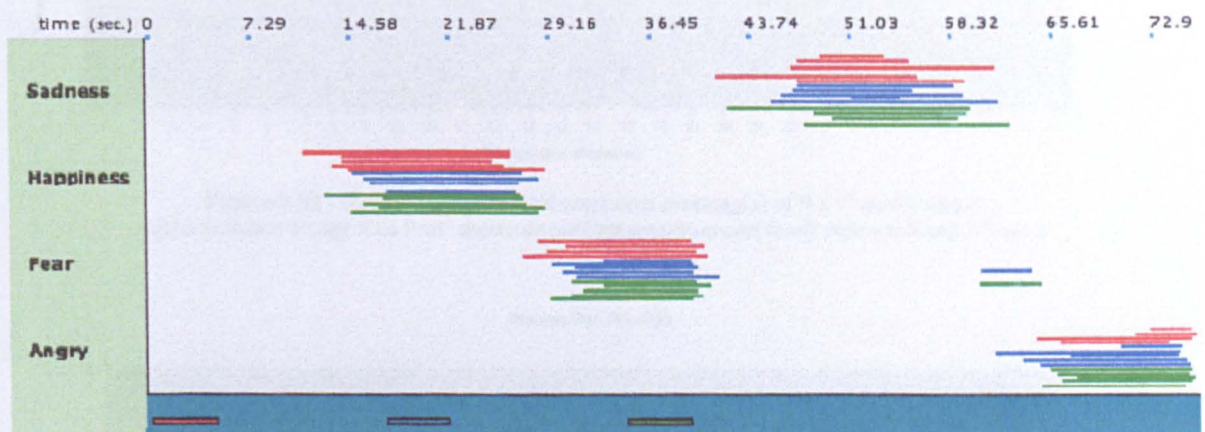
Figure XX - Profiles of the perceived emotional meaning(s) of the 2nd performance of the Schubert's Lied 'Die Post'

The two previous graphs (Figure XIX and XX) show the final results of the first and second performances of the Schubert's song, 'Die Post'. They were printed out immediately after each performance and worked as the singer's feedback of the audience perception and recognition of the performed emotional meanings. The graphs produced present the first part of the analysis here, and the questionnaires are discussed later.

The next two graphs (Figure XXI and XXII) represent the final results summarising with more precision the audience's emotional perception and recognition in Experiment-IV, these are the data to be analysed and discussed:

First Run through

Figure XXI - Graph of the perceived emotional meaning(s) of the 1st performance
of the Schubert's Lied 'Die Post' shown as performance moments (each moment being 2.5 sec.).

Second Run Through
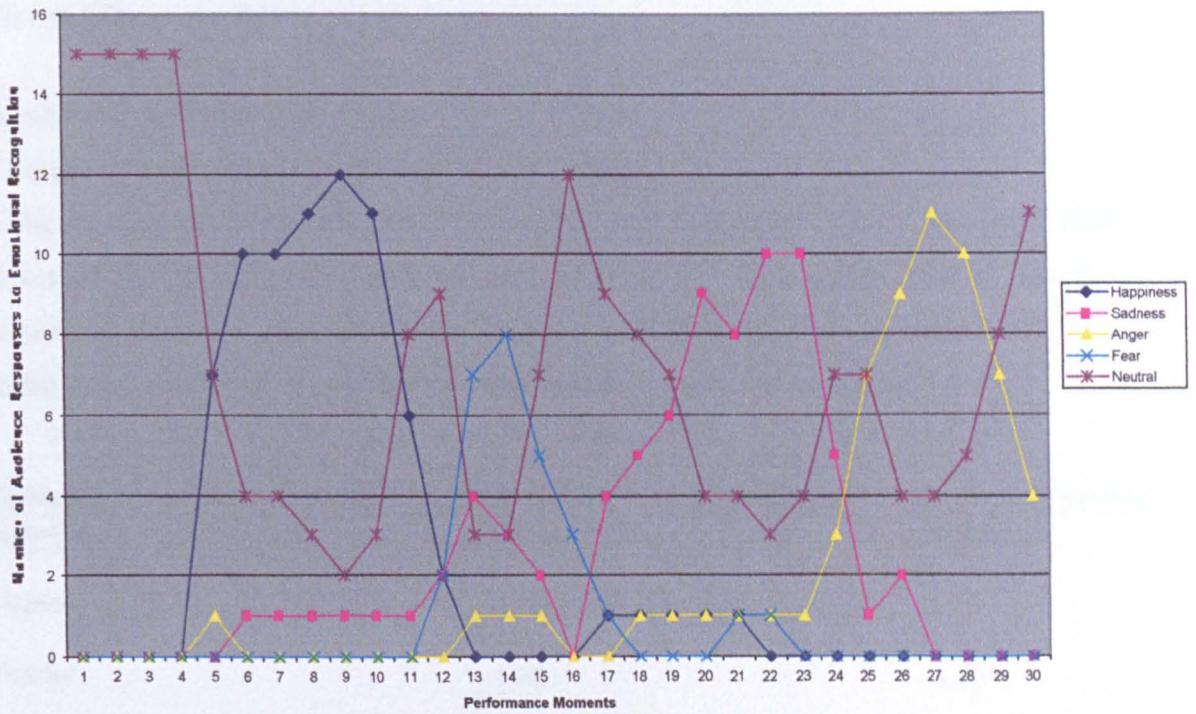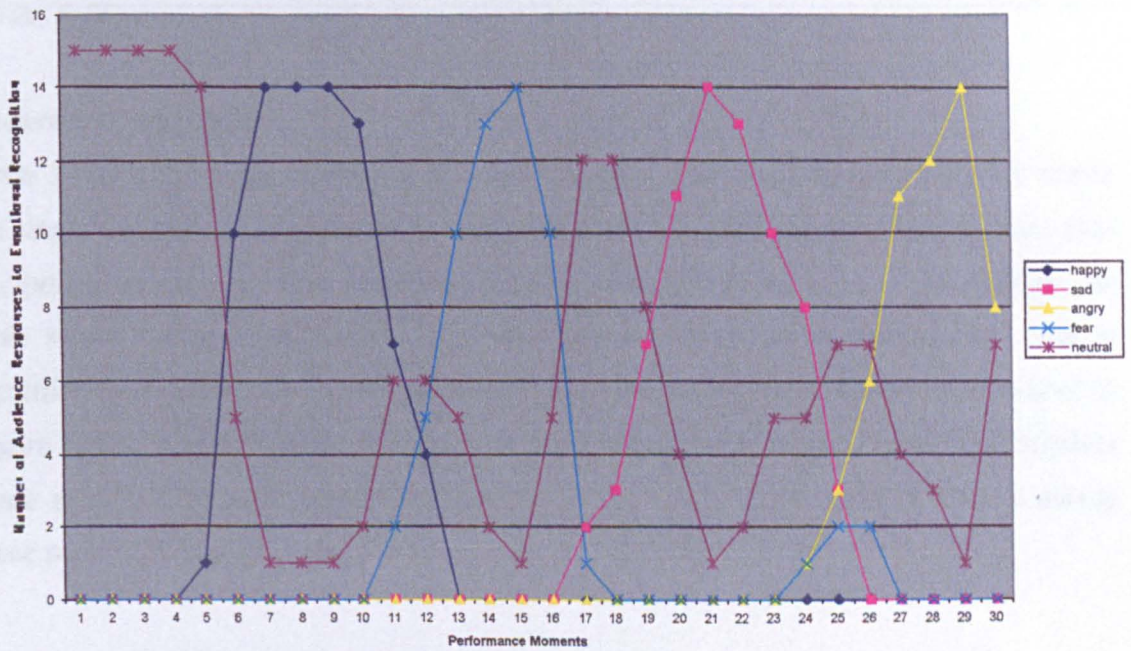

Figure XXII - Graph of the perceived emotional meaning(s) of the 2nd performance
of the Schubert's Lied 'Die Post' shown as performance moments (each moment being 2.5 sec.).

With these data in hand, a series of statistical tests were undertaken to try to understand the data as fully as possible.

Analysis of real time experiment

Comparison of overall distribution of key presses.

Table IX shows that from the first to the second run through participants rated significantly fewer of the 30 moments (each moment being of 2.5 seconds) as neutral and rated significantly more as representing fear (t-test, $p < 0.05$). There were no overall differences in the mean number of presses for the other emotions according to a t-test.

Table IX: Mean number of moments (of 30x2,5 sec. excerpts) recorded to indicate each emotion

Emotion	1st run through	2nd Run through
Happy	4.93	5.26
Fearful	1.86	4.46*
Sad	5.10	5.10
Angry	4.06	3.40
Neutral	14.0	11.6*

* $P < 0.05$ (t-test)

Accuracy of responses

Table X gives the mean number of the audience identified moments that correctly match the singer's intended interpretation during peak periods of each emotional expression. This is reported for each emotion separately and also an overall score is given. In addition, the table shows the average number of neutral presses given during this period. Fear is identified for significantly more time during the 4 emotional peak moments (as indicated in Figure XXIII, pg. 198) in run through two, than in run through one. Overall, participants made significantly more correct key presses in run through two rather than one during these peak emotional periods.

These errors in 'Run-through One' are accounted for mainly by neutral key presses - that is, there were significantly more neutral presses in 'Run-through One' than 'Run-through Two' during the peak emotional periods. The total of correct responses, plus the neutral errors account for all 30 moments. While there was a suggestion that in 'Run-through One' fear was more easily confused with anger than in 'Run-through two' this was not confused for significantly more key moments in 'Run-through one' than 'Run-through two'.

Table X: Mean rating of the correct responses during the peak emotional moments

Emotion (key moments)	1st run through	2nd Run through
Happy (8-9)	2.93	3.66
Fearful (14-15)	1.53	3.00*
Sad (21-22)	2.46	3.13
Angry (28-29)	2.13	2.86
Overall correct (all 30 moments)	9.01	12.67**
Neutral presses (all 30 moments)	5.33	2.33**

* p<0.05 (t-test)
** p <0.001 (t-test)

DIE POST *Franz Schubert*

		Singer's intended peak of the 1st expressed Emotion Happy (1) (~20,00'-22,50')						Singer's intended peak of the 2nd expressed Emotion Fear (4) (~35,00'-37,50')						Singer's intended peak of the 3rd expressed Emotion Sad (2) (~50,00'-52,50')								Singer's intended peak of the 4th expressed Emotion Angry (3) (~70,00'-72,50')			
		1st moment	2nd moment	3rd moment	4th moment	5th moment	6th moment	7th moment	8th moment	9th moment	10th moment	11th moment	12th moment	13th moment	14th moment	15th moment	16th moment	17th moment	18th moment	19th moment	20th moment	21st moment	22nd moment	23rd moment	24th moment	25th moment	26th moment	27th moment	28th moment	29th moment	30th moment
		Answer RT 00,00-02,50	Answer RT 02,50-05,00	Answer RT 05,00-07,50	Answer RT 07,50-10,00	Answer RT 10,00-12,50	Answer RT 12,50-15,00	Answer RT 15,00-17,50	Answer RT 17,50-20,00	Answer RT 20,00-22,50	Answer RT 22,50-25,00	Answer RT 25,00-27,50	Answer RT 27,50-30,00	Answer RT 30,00-32,50	Answer RT 32,50-35,00	Answer RT 35,00-37,50	Answer RT 37,50-40,00	Answer RT 40,00-42,50	Answer RT 42,50-45,00	Answer RT 45,00-47,50	Answer RT 47,50-50,00	Answer RT 50,00-52,50	Answer RT 52,50-55,00	Answer RT 55,00-57,50	Answer RT 57,50-60,00	Answer RT 60,00-62,50	Answer RT 62,50-65,00	Answer RT 65,00-67,50	Answer RT 67,50-70,00	Answer RT 70,00-72,50	Answer RT 72,50-75,00

FIGURE XXIII: Some of the music structural features selected by the singer to underline and convey the strongest interpretative moments of the performance of the Schubert song, where the intended emotional meaning communicated aimed for the audience's recognition of a categorical basic emotion.

Triangulation Data

Audience Interview Responses

In order to explore further how the audience regarded the test and how they made their decisions, a face-to-face interview was undertaken with each member (see Method and Appendix C for further details). These were based on Smith's (1995) Interpretative Phenomenological Analysis questionnaire design and analysis technique. The following sections present the key themes emerging from the questioning and are presented in terms of the numbers of respondents making the discussion point.[47]

Clarity of cues

Run-through One versus Run-through Two

Without exception, all respondents found it easier to judge the second interpretation. For example, two interviewees made the following points:

'Well, when I was viewing the first performance I would have to think : What's he doing here? I'd have to really think about what he was presenting to me. On the second version, it seemed a lot clearer. I knew, in particular that there was scary stuff going on!'

'The second performance seemed more emotionally clear. The face and voice looked/sounded more direct. May be it was me getting used to the task, but I don't think so. The second one was clearer, for sure!'

It was important to consider whether or not previous experience had an impact on the audience's perception of the expressive content, but in fact, these data showed no differences. The two examples above come from a musician and a non-musician. Similarly, men and women seemed to be equally able to recognise the emotional cues, though sometimes the vocabulary used by them was different. For example:

'Emotionally charged and full of passion!' (Female)

'I found the performance to be very expressive, on the whole, it was convincing'. (Male)

---

[47] In this discussion, two examples are given as indicative samples of the responses obtained.

Thus, an interpretation that could be made is that females are more ready to use expressive language in talking about emotional matters. But, this was not a principal focus of the study, and so will not be explored further here. However, it is worth noting that Toni is of course male. And, if we look back to Experiment-III we discover that in fact he was more able to present the emotional content of his songs to the audience than Silvia. Thus, it would be unwise to make any sort of generalised statement about gender and emotional expression on the basis of the current data.

Types of cues

Here, it seemed that the stronger the facial gesture (one could say the more clearly marked) the better:

> *'When he was smiling with such intensity, I could feel the happiness bouncing from him. I could also feel myself sort of smiling too.'*

> *'The cues were quite clearly defined, but things like fear were certainly clearer in the second performance. Somehow, there was more definition. I felt it in my face too.'*

We must not confuse clarity of definition with exaggeration, however. In the practice trials for this experiment it was noted that when too much gesturing – vocal or physical- was made, the performance seemed like a caricature in its over-exaggeration. So, authenticity may have a key role here: clear, but somehow truly felt. Indeed, in the following section in which the performer's own thoughts are discussed, authenticity becomes a major issue.

Familiarity, liking and accuracy

Examining the responses, it became evident that there was some sort of correlation between familiarity and liking and the more the piece was liked, it was generally familiar to the audience members:

> *'I don't know this kind of music. Emotions were clear, but it is not the sort of thing I would say I liked. I can appreciate the emotional content, even feel some of it myself, but I don't really like the style'* (Non-musician).

> *'I simply love the way Schubert is able to fuse structural effects and raw human emotions. Toni conveyed this really well. Schubert sung well, in my opinion!'* (Musician).

The two quotes also provide evidence that accuracy of emotional perception was not influenced in this experiment at least by liking. All audience members were fairly competent in their judgements, certainly second time around.

Eyes and ears

The audience members seemed to rely on both oral and visual cues, but visual information such as widening of the eyes, or a sudden movement and consequent change of expression was easier to use.

> *'He was certainly most expressive with his face – speaking to me with the eyes sort of thing'.*

> *'Facial cues were very moving!'*

The audience strategies to recognise an emotional state were based largely on feelings the singer 'projected' into them:

> *'I could feel his anger inside me'.*

> *'The last moment of the song was so strong that I even felt scared!'*

> *'Those moments when I had a shiver down my spine reminded me of childhood nightmares'.*

So, it seemed that previous and personal experience of an emotional state enabled the perceiver to empathise or even experience a similar emotion to that of the performer. Sometimes, even, a kind of confusion seems to appear between perceived and felt emotional meaning. Finally, it is worth noting that although many commented on the

amount of concentration required to undertake the task, none of them found the software of the task itself difficult to understand or control. At this point, Toni's own reflections can shed further light on the results.

Toni's self-report

Of the approach he adopted, Toni said:

"My approach was a sort of go inside myself to put out the emotion. It was a kind of a self-induced emotion. A recollection from other emotional states or moods... a kind of empathy for emotions is something... Sort of feeling the emotion... Yes, I would say it is a sensitivity to the genuine emotion! So for instance, I am not enraged, but I have all the physical-mental attributes of anger present. So during the performance, for instance, I felt that as though what was being projected visually and acoustically is what I felt inside. The access to the audience's feedback and to the videotaped performances allowed for a better self-control of the facial and vocal expressive cues second time, I know it did! And therefore, I could feel that the second performance really matched better the intended emotional expressiveness. Nevertheless, I think in all cases the emotions are clear... Where I do repeat the emotions, sometimes one version is visually slightly clearer than another, but the reality of the emotion is always present. It is more a question of style of presentation rather than genuineness of content. Of course, I always feel that when the emotion is more authentically felt the audience gets better the meaning of what is being expressed. I can also see that I do certain things physically to accommodate to the technical aspects of voice production. It does not detract from the emotion, but it is an element through which my performance is mediated. For instance, in anger, I can see that I am really concentrating on keeping my resonating cavities open and focus on my voice high into the upper part of my cheeks. Because the face moves intuitively for an emotion (even if I can correct and make clear the facial expressive gestures by learning how they best express the different emotions) I don't feel that the general aspects of the singing technique really interfere with the expressiveness of the facial muscles. Of course, in some extreme cases of emotional expression, such as anger, for instance, the singer should be aware of specific muscles, such as the Platysma, whose action could interfere with the required freedom in the singing vocal technique. So, it seems that it was a good and potentially very useful practice to experiment with the emotions to explore which intentions are best and how they interact with the technical facial position...To get the musculature at its optimum level. I think this is a very useful approach for feedback for change, but also to affirm and give confidence to the performer. Also, I could, at points, especially in happiness and sadness, notice the audience engaging facially in my gestures: we were sort of sharing expressions."

So, it becomes apparent that the technique is aiding Toni too. He made no comment about the procedure involved in watching the video and the audience responses between performances, but it can be assumed that 3 minutes of concentration was an acceptable and useful period of time for such a focus.

Discussion

By simply observing the two graphs (Figures XXI and XXII) it is apparent that the second version reveals a more defined performance of the different emotions. This is then particularly well clarified in a combined musical-audience response-visual display presentation of the results created in Figure XXIII. In summary, these data indicate that the second performance has been a more successful version of the same intended interpretation of the Schubert song. Table X shows that performance two presented better stimuli for the judgement of emotional meaning. The audience answers presented a higher mean rating of accuracy for all the communicated emotional meanings within the second performance of the song. This means that the recognised emotions match the analysed content of the song better in the second performance than in the first performance. However, some confusion still seems to appear in the recognition of the last peak of emotional meaning communicated in the second performance. In fact, a few responses of the audience to the 'Second run through' still indicate fear instead of anger, as was intended in Toni's interpretation. This might be explained by the confusion between 'perceived' and 'felt' emotion (see final discussion in Chapter 5). Of course, it can be suggested that every time a performance happens for the second time it becomes clearer and more meaningful for a perceiver than the first time. But, the preparation and development of the Experiment-IV has revealed that this is, most of the time, not at all the case, and that a successful performance depends really on the ability of the performer to express clearly the meaning he/she intends to communicate. In fact, in the development of the Experiment-IV, it was verified that a second version of the same interpretation sometimes generated, through an 'overdoing' of the emotional cues an excess of emotional information, generating a less

clear profile for the audience's responses. So, the conclusion might be that it is with the rehearsal and study of the different videotaped performances and the results of the feedback from the audience's responses for each performed version of the same song, that a more accurate performance can be effectively achieved. It certainly provides the performer with information about which facial and vocal cues might be used in order to be able to explore the personal expressive characteristics for an accurate emotional meaning communication.

When the Experiment-IV was concluded, an interview with all members of the audience was undertaken (see Appendix D). These data suggested and corroborated the clearer profile of the second performance. The general opinion was that the facial and vocal cues were performed clearer and more accurately in the 'Second run through' of the song. The singer, himself, was more convinced of his interpretation of the second version and, therefore, felt that the expressive moments of the second performance were communicating the emotions with more authenticity. It seems that the audience found the task tolerable, in that they all did it without problems. Furthermore, musical experience did not affect the ability to detect the emotions, though visual cues were more easily judged. This finding corroborates Davidson (1993), who found that vision was indeed a clearer indicator of expression in musical performance. Moreover, Toni's own comments and those of the audience show that the projections and empathy are strongly identified and used as the perceptual means of 'understanding' the performance. A psychodynamic interpretation of these data is possible. Perhaps a link to the previous studies by Kristeva (1989) and Bierwisch (1979) is most appropriate here: the audience experience the emotion themselves, and so does Toni. But, in both cases they are not 'really' angry or sad, rather they are recognising and identifying with those emotions. Rather intriguingly, the degree of 'shared' experience seemed to include a transfer of cue too – both smiled. This might be linked to conversational exchanges, even mother-child mirroring as referred to by Gallese & Goldman's (1998) and Trevarthen (1999).

So, the questions left opened at the conclusion of Chapter 3 (such as, how is the performer-audience emotional communication in the performance of singing under realistic conditions? Is the emotional meaning accurately perceived and recognised in a real life

performance as it is in the laboratory conditions? Is the emotional meaning of the music structure consistently vocally and facially communicated when performing a full song under the realistic conditions of a concert room? How reliable are the findings of Experiment-III when submitted to an ecological valid investigation?), seem to get a clear answer resulting from the analysis and results from Experiment-IV. In fact, Experiment-IV seems to reveal a high level of performance consistency with the musical structure, and the audience perception and recognition of the emotional meaning communicated seems to be accurately matching the singer's interpretation during intentional peak periods of emotional expressiveness. So, by comparing the analysis of both Experiments (III and IV) it seems possible to conclude that the findings of Experiment-III are consistently supported by the ecological validity of the results of Experiment-IV. It is also noting how the interface between music structural features and expressive vocal and facial cues really aided to clarify the perception and recognition of the emotional meaning expressed. In fact, as it can be seen in Figure XXIII, some of the audience responses to emotional meaning during the "First and Second run through" (especially happiness and sadness) appeared even before Toni would start to sing. This seems to indicate that the musical structure was in itself able to generate the perception and recognition of the emotional meaning that Toni, moments later, would be corroborating through the facial and vocal emotional expressive cues of his performance. This could, for instance explain, why happiness and sadness were the best recognised emotional meanings in the 'Die Post' performance, while in the 'Erlkönig' (Experiment-III) the sung phrase the best recognised in emotional meaning was anger. That is the musical structure features used for happiness and sadness are really aiding strongly and clearly to strengthen the emotional meaning communication of Toni's interpretation of the song. In Experiment-III, on the contrary, the easiest Emotion to be identified by the audience was anger. Given the method and the experimental procedure used in Experiment-III (a laboratory with no musical context), this result might be explained by the fact that the emotional meaning perception and recognition was dependent only on the singer's expressive cues. Thus, it might be concluded that when the singer's facial and vocal expressive cues corroborate intentionally the meaning of the selected expressive features of the musical structure, the accuracy of the audience in perceiving and recognising the emotional meaning of the music performed is much higher than in the remaining parts of the song. Interestingly, it seems that the perception and

recognition of emotion in this experiment is mainly concentrated around the intended emotional peaks of the performance. This does not mean, however, that there is no emotional meaning communicated in the rest of the song, but only that the moments of higher level of audience's accuracy are connected with the strongest emotional content of the music performance and the capacity of the singer to communicate discrete emotional meaning through the facial and vocal expressivity in singing. Intriguingly the singers in interview did not focus on this point, rather on the fact that emotion was a key part of their experience of the song, whether on-going or perceived in 'phrases'.

A further discussion on the successes and limitations of the techniques used in Experiment-IV can be found in final conclusions, in Chapter 5.

CHAPTER 5

5. CONCLUSIONS ON THE EMPIRICAL EVIDENCE ON SINGING AND EMOTION

THE RELEVANCE OF THE EXPRESSION OF E-MOTION IN THE FACE AND VOICE OF THE SINGER IN THE COMMUNICATION OF EMOTIONAL MEANING IN SINGING. A "PERFORMER'S EXPRESSIVE TOOL".

5.1 SUMMARY OF THE EMPIRICAL STUDIES

The results of the previous chapter indicate and corroborate former empirical findings that the communication of emotional meaning in singing is directly related to the singer's intention and capacity to translate the emotional musical meaning through clear acoustical (vocal) and visual (facial) expressive cues. The performer's interpretation, an expressive intention that should try to corroborate the emotional meaning revealed and conveyed through the musical structure, should find in these expressive cues an important creative device at the service of the music and its meaning. The performance could be seen as a relevant and important way to convey the musical meaning and a structured medium to express and communicate its emotional content, with movement corroborating the music and that in turn corroborating the movement. Of course, these expressive devices, movements and resulting sounds, do belong to our everyday vocabulary of emotional communication. They function at the primary level of expressive intentional movement, which has been labelled as the 'Indexical-qualitative coding' of music performance (see in Chapter 3, pg. 108). But, if the performance of music is supposedly to be studied as a communication process, these basic expressive emotional mechanisms should be considered of primary importance, precisely because of their innate and universal nature. Music is nowadays accepted as a part of everyday listening and perceiving experience (Small, 1987). Thus it seems natural to accept that music expressiveness is related to the movements and gestures of everyday communication. According to Swanwick (1988:64) the expressive character of music is a direct development of first childhood mimetic representation. Of course, the performer's movements are also related to the technical adjustments for ideal tone production, and to the more structural and syntactic components of musical meaning. Nevertheless, and because the primary aim of this study is the communication of emotional meaning in music performance, the empirical investigation undertaken in previous chapter focused more on the expressive cues that could help the

musician (here, the singer), or a music student, to improve expressively his/her capacity to communicate an intentional emotional musical meaning.

5.2   THEORETICAL IMPLICATIONS

(Overview of the empirical evidence revealed in this Thesis: Innovation and Novelty)

The empirical evidence found in the experiments in Chapter 4 reveals the importance that the expression of emotion in the face and the voice of the singer have in the communication of the emotional meaning within the performance of singing. It also suggests that, though voice and face together provide the most accurate stimulus presentation for the judgement of emotion in singing, facial expression provided clearer information about emotional meaning than vocal expression alone. It also reveals that, the facial expression of emotional meaning in singing involves the same type of musculature involved in everyday facial emotional expressiveness (at least for the four basic emotions here investigated). These facial expressions are distinct and quantitatively measurable. Also, whatever the degree of authenticity of a certain emotional expression the musculature involved is the same, only with less degree of activity. Nevertheless, the degree of accuracy involved in the perception of emotional meaning by the audience will rise, according to the degree of authenticity achieved by the performer. Another important conclusion to be drawn from the empirical evidence found in the experiments in Chapter 4 reveals that a Non Verbal mode of response is a more accurate way of labelling the expression of emotions in music performance. In this sense, several theoretical implications can be drawn from these results, and these are discussed below.

5.2.1 THE PERFORMER (SINGER) AND THE MUSIC/PERFORMANCE (SINGING) AS A PART OF A COMMUNICATION SYSTEM

One of these implications is the fact that music might be seen as a part of a communication system where the performer might be viewed as an interacting pivot between several factors: the music (the composer), the listener, and the context. The interactive role of the performer as a part of a communication system in which the "composers code musical ideas in notation, performers re-code from the notation to musical signal, and listeners re-code from the acoustic signal to ideas" (Kendall & Carterette, 1990) seems to assign to the performer's movements, intentionally used to produce musical interpretation, a crucial important role in the process of generating and communicating musical meaning.

According to Davidson (2002:242), "the bodily production of music involves general motion and specific localised culturally learned gestures for technical, musical, and socially communicative purposes." So, it would appear that beyond production, and from a perceptual perspective, body movements would assist the co-ordination between co-performers and aid the audience comprehension of the musical performance (Davidson, 2002). This is exactly what Experiments-III and IV seem to reveal. They show how the performer's movements and gestures, either vocally or facially, corroborate the intended meaning of the musical structure, and also how they add an important amount of information to help the audience to understand the full message of the communication process. Furthermore, this seems to be in agreement with Gibson's claim that meaning, however universal or fine-spun it might be, is clearly dependent on sensitive stimuli:

> "No symbol exists except as it is realised in sound, projected light, mechanical contact, or the like."
>
> (Gibson, 1966:26)

The sound and gestures produced by and according to the performer's interpretation of the musical structure are of course the communicative content in a music performance, which will be then de-coded by the listener and perceived as the musically intended or embodied meaning. So, one of the main contributions of the empirical evidence found in the experiments undertaken in this thesis seems connected to the fact that music in order to have access to its ontological status must be 'played'. In other words, no music exists if it is not 'performed music', i.e., if it is not the result of an embodied act of musical meaning.[48]

---

[48] I see music in the 'score-state' as a kind of 'musical virus', a not-actually-developed 'musical idea'. In this state of latency, music waits for a body, or a living system, where it could take place and existence as a musical event within a process of communication. In order to be able to occupy a time and a space in the everyday world and thus be subjected to the physical, psychological, and social constrains of the "real world", music needs to be performed.

5.2.2 THE MUSIC/PERFORMANCE (SINGING) AND THE ECOLOGICAL LEVEL OF MUSICAL AWARENESS

Music as an ecological event, as explained in Chapter 3, might be connected with different levels of awareness. As Baily (1985:258) suggested, there are multiple modes, or levels, of musical awareness:

> "Music may be as much a motor event, as a sonic event, as well as, of course, a social fact."
>
> (Baily, 1985:258)

Traditionally and especially connected to the last three centuries of Western musical meaning and value, music has been conceived mostly as an autonomous event. This is, however, according to contemporary authors, just "one mode of musical awareness among several" (Shove & Repp, 1995:59). According to this level of awareness, the movements of the performer only exist outside the 'music proper', and therefore the autonomous level creates an ontological gap between music as heard and music as performed. At this level, movement in music is only considered abstractly, virtually or as an illusion. The results of the empirical studies of this thesis, however, show otherwise. According to the ecological claim, it would seem that musical movement is above all directly connected to the performer's movements. Music is considered as existing factually in the world, and the meanings conveyed through the sounds of music belong, one way or the other, to a human communication process. For the ecological approach, as Clarke (2001) states, the source of sound "is specified in the stimulus information as it arrives at the sensory system of a perceiver: its spatial location, material, mode of excitation, separation from other sources, and so on" (Clarke, 2001:219). Therefore, and because the focus of this thesis has been the emotional character of music, the empirical studies undertaken have considered the gestures and movements of the performer as being consistent with the emotional meaning conveyed and perceived through the performed music. The intended emotion has been expressed, i.e., specified in the stimulus information of the performed music and communicated (visually and acoustically) to the sensory system of the perceiver, who then, based on this information, has tried to recognise the intended emotional meaning.

A main conclusion of Experiment-III was that the vocal and facial expressions used by singers when trying to communicate the emotional meaning found in the musical structure

were similar to the vocal and facial expressions used for emotional expression in the everyday world. Thus, whether intrinsically or extrinsically structurally coded, the emotional musical meaning comes to the performance level as an act of embodied meaning and it is constituted by the same bodily movements and gestures perceptually available and comprehensible to audiences as well as to the everyday emotional meaning communication. It seems possible to conclude that, in a musical performance context, the emotions portrayed by the singer/performer, though performed emotions, are not 'faked' or abstract or illusory emotions. Their content and form are built on real emotions, in the sense that they are really felt as recollections of these emotions, and that their expressive characteristics are similar to those from real emotions. (Of course, performers do not have to feel an emotion to convey it to an audience, but, according to the empirical evidence revealed in this study, feeling the intended emotion will intensify its expression and authenticity, and is likely to make it more perceptible to an audience. Finally, it will be this fact that differentiates a good from a bad performance, what will make an interpretation unique and will introduce in a performance the 'intangible' element of *artistry*). The musical performance artistic domain is thus not a domain autonomous and separated from the real world and the real gestures and emotions, but a domain that takes its authenticity from the ecological reality – and here I choose to substitute the word validity - of the music expressive event. The emotional meaning of a musical performance is comprised of all those movements and patterns of movement whose general characteristics are similar to bodily movement symptomatic of human emotions and that any performer should be able to express through his/her bodily gestures. As Shove & Repp (1995:60) state: "On the ecological level the source of movement is the human performer. In this regard, musical movement is human movement". Of course, to feel a certain emotion is not to express it. Performance expressiveness is not achieved only through musical and technical expertise. The performer must also be taught how to express feelings and emotions whilst performing. This 'know-how' has to be learned and rehearsed in order to become natural, intentional and expressive. Based on these arguments, the power of the Experiment-IV was to introduce an empirical tool that could assist the singer to improve his/her capacities to express emotional musical meaning (see under, 5.3.2 'A performance practical application').

5.2.3 Discreteness and accuracy in emotional meaning communication in singing

The main focus of the experimental work of this thesis investigated if the performer's capacity to communicate discrete emotional meaning in singing was accurately perceived by an audience. The experiments undertaken have also explored whether the emotional meaning communicated by the performer's body movements was simultaneously enhancing the inherent emotional meaning implied by the musical structure and/or by the musical context. In fact, the results indicate that the communication process, though enhancing the emotional content of the musical structure, was mainly based on the singer's ability to express facially or/and vocally the emotional meaning of the musical structure within the performance of singing. Nevertheless, a careful reading of the graphs of Experiment IV (see Figures XIX and XX, pg. 192 and 194) shows that some of the audience's categorical responses were already present long before the planned moment for the emotional peak, and sometimes even before the singer started to perform. In fact, it seems that the audience was able to recognise some of the intended emotional meanings just by hearing the music structure played in the piano (it is the case, for instance, of happiness – the staccato articulation of E flat major - and sadness – the movement into the minor mode). This fact suggests that, the interpretation of a musical piece should not be done against the music structure (or the composer). It seems thus relevant to underline that, if the communication process in music performance is, as stated, primarily dependent on the performer's ability to express the music's emotional meaning, the role of the performer acquires its really artistic strength by corroborating and aiding to clarify the emotional meaning conveyed through the musical structure. So, the performer should be aware that, by being able to adjust the meaning conveyed by his/her performance to the emotional content of the musical piece, his/her performance will gain in consistency and, therefore, in expressiveness.

Based on the Brunswick's (1956) 'lens model', Juslin (2001) named the ensemble of expressive cues that aid to convey the emotional meaning of a musical piece (for instance in case of sadness: slow mean tempo, legato articulation, low sound level, dull timbre, slow tone attacks, etc) as the *probabilistic redundant musical expressive cues for emotional communication*. These cues should be seen as an important performer's expressive means to convey the emotional content of the musical structure. Moreover, the singer should be aware that, besides these expressive musical cues, he has at his/her service other means of

expressiveness (mainly, the vocal and facial non-verbal communication) which might strongly corroborate the acoustical features of the musical structure. Vocally, for example, a slow tempo of vocal articulation, a round vocal envelope, a low level of harmonics, soft dynamics, a slow attack, are some of the vocal cues that could corroborate and underline the above mentioned musical expressive cues for the communication of *sadness*. Indeed, evidence from facial electromyography and self-reports to measure emotions claim that we do indeed react emotionally to the performance of music as we do to vocal expressions (Lundqvist et al., 2000; Witvliet et al., 1998). And, Scherer (1995), Scherer & Siegwart (1995), and Banse & Scherer (1996), among others, have shown that vocal expression is a strong medium for emotional meaning communication. However, music performance, especially in case of singing, also depends on visual information in the process of meaning communication. In fact, empirical evidence on facial expression by Ekman & Friesen (1976) has shown that facial expression is a strong medium for emotional communication. Similarly, the experiments of this thesis suggest that, at least for the four categories of emotional meaning here investigated, facial and vocal non-verbal expressiveness worked as strong allies in the communication of the emotional meaning conveyed through the musical structure. Within the different musical contexts of the experiments undertaken in this study, the results showed that the specification and accuracy of the emotional meaning communicated was always related to the performer's capacity to perform (facially or/and vocally) the intended emotional meaning in a categorical way. Also, the conclusions of Experiment-III clearly pointed to a sort of hierarchy of distribution of the different expressive means and of their combinations. So, the clear and strongest emotional meaning effect was connected to the combination of a categorical facial and a categorical vocal expression. The second strongest medium for emotional communication was the use of a categorical facial expression. The third strongest medium of emotional meaning was the presentation of a categorical meaning when vocally expressed. All these means of emotional communication are of course especially strong if they are connected with and supporting a musical structure whose expressive features (intrinsic or extrinsic) are clearly consistent with the emotional meaning facially or/and vocally expressed. The high degree of accuracy in the recognition of the different emotional meanings in singing suggests a high degree of consistency between the musical structure content and the performer's expressiveness, and that the perception of emotions in singing is natural and effortless for the large majority of the perceivers, thus confirming the studies of Scherer & Siegwart (1995) presented in the Literature Review. Turning briefly to the audience and their results, it is evident that the concept of emotional transfer and the related clear detection of cues

were highly significant perceptual achievements. Audience gender and musical experience perhaps led to a focus on different elements of the performance overall, but all audience members seemed to be able to detect the emotional content through a sort of empathy process. The singer too used this means of experiencing the emotions when he modified his interpretation between 'Run though One and Two' in Experiment-IV. Linking back to the theories exposed at the start of this thesis (Hatfield et al., 1995) it is possible to conclude that a kind of 'empathy-authenticity' process was at work. In fact, Juslin (2001) claims that "it is possible that the expression of music performance may arouse emotions in listeners through a process of so-called emotional contagion." This sort of process might equally be involved in the process of emotional perception and recognition through a sort of 'motor mimicry' or 'mimetic gesture' (see Cox, 2001). Lipps (1909), for instance, argued for an understanding of the emotions based on the human capacity of mimicking the expressive gestures seen in the others. Thus, one might conclude that, singing may have a major contribution to make in the clarification and improvement of the emotional perception and experience in music, because the singer has access to some of the most relevant means in emotional communication in musical performance: the face, the voice, the music, and the text. It could be that the singer not only uses his/her own experiences and feedback, but that of the audience too. Indeed, looking into the audience Toni could see empathic facial gestures related to the most expressive moments of the performance, especially in the happy and sad emotional content. An interesting result of Experiment-III corroborating this idea was that Non Verbal recognition of emotions was higher than verbal recognition, probably as a consequence of the mimetic identification with the portrayed face made by the perceiver. In summary, the empirical work has clarified the theoretical position argued in the first three chapters of this thesis. It was linked in the analysis of *Die Post* but can be re-evaluated in the following paragraphs:

Indexical-qualitative coding: Of course, and as referred to before, the empirical investigation in the scope of this thesis, undertaken in Chapter-4, is mainly concerned with the facial and vocal expressive elements of the music performance emotional communication which are also a part of the emotional expression in everyday communication (see Appendix F for different examples).

Iconic-structural/dynamic coding: As it is also referred to in Appendix F, some of the emotional expressive vocal and facial elements may also be enhancing some of the music's

structural features identified in Chapter-3 as belonging to the Iconic-structural/dynamic (introversive/extroversive) coding. For example, when voice and face expressiveness enhance the emotional meaning of the music, structural features in bars 12 to 16 in the song *Die Post*, in Experiment-IV (Iconic-structural-extroversive coding) or when in the same experiment the vocal and facial emotional expressive elements enhance the emotional content of the music structural features in bars 37 to 46 (Iconic-structural-introversive coding). Another example, where it is possible to observe the introversive and the extroversive structural dynamic coding at work simultaneously can be seen in bars 27 to 36 with the movement to the minor mode structurally confirming the emotional transition to sadness (Hevner, 1936; Rigg, 1937; Kleinen, 1968; Wedin, 1972; Crowder, 1985; Krumhansl, 1997; Peretz et al. 1998) - *Iconic-structural-introversive coding* (see Appendix F, pg. 248/49) – and, for instance, the descending minor third followed by an ascending minor second, in bars 34/35, working musically as a mimetic movement of a sigh, underlining the melancholic feeling of the traveller (see Cooke, 59; Maher & Berline, 1982) – *Iconic-structural-extroversive coding* (see Appendix F, pg. 245/46). Another possibility to observe the work of the Iconic-structural-extroversive coding can be seen in the way Schubert in the beginning of Die Post musically describes the arrival of the mail coach and the excitement of the traveller by using the lively 6/8 rhythm with the piano staccato articulation together with a light and bright musical structure. The melodic line brings with the interval of the 3rd major the indexical structural element which points to the announcement of the postman's horn call (see Appendix F, for other examples). All these introversive and extroversive music structural features can be enhanced by vocal and facial expressive elements which may then strengthen the communication of the intended emotional expression. For instance, in the example of sadness, the minor mode can be enhanced by an expressive legato articulation (Juslin, 1997) on the vocal part (contrasting with the staccato articulation for the opening excited and happy atmosphere of the beginning song) corroborating the feeling of sorrow together with a low sound level (Juslin, 2000), slow tone attacks (Kotlyar & Morozov, 1976), or even a dared flat micro-intonation (Baroni & Finarelli, 1994). Facially, the expression corroborating the music emotional content of sadness could be enhanced through the activity of the muscles Corrugator, Triangularis, Depressor Supercilii, Mentalis and Depressor Labii Inferioris (see Figure XVII and Appendix F for other examples, too.)

Symbolic-associative coding: Of course, this level of musical emotional communication, though indirectly connected to the articulatory and interactive movements of the performer and to certain patterns and specific dynamic characteristics of the musical structure and transformational features of the performed music, is mainly based and generated on occasional and idiosyncratic factors associated with the emotional life of an individual's or collective's memory (Scherer & Zentner, 2001). Thus, though able to generate individual and collectively important emotional meanings, this level of music emotional coding is based on arbitrary or contingent elements and not in a necessary relationship between the inherent characteristics of the sound or generating movement and the characteristics of the phenomenon evoked or referred to, and so have not been the focus of this investigation.

5.3   PRACTICAL IMPLICATIONS

5.3.1 ESTABLISHING AN ANALYTICAL FRAMEWORK FOR MUSIC PERFORMANCE: AN EXPRESSIVE PERFORMANCE TOOL FOR SINGERS AND SINGING TEACHERS.

The empirical work and evidence of this thesis suggest that within the signifying process of singing, facial and vocal behaviour play an important role when communicating and interacting with an audience. Thus, it seems that the study of singing, as both a creative and a meaningful activity, would require an appropriate set of expressive cues, which need to be identified by the singer and brought to a conscious level in order to fully explore the expressiveness of their characteristics and of their meaning. Through the results of the empirical study of this thesis, a contribution to a higher level of self-perception within the performance of singing and, consequently, to a higher level of musical performance and musical meaning expression in singing has been sought (see above, Experiment III and IV). The intention of the empirical research undertaken has been to study and to improve the capacity of a singer to express and communicate emotional meaning to an audience within a music performance. In this sense, an attempt was made, especially through Experiment-IV (see above explanation), to design a means of developing the singer's expressive potential. Software was elaborated in order to register the perceivers' recognition of the performed emotional content and identification timing (Appendix G).

The 'expressive performance tool' was thus elaborated in order to check and improve the expressiveness of the singer. Of course, in order to obtain a high degree of authenticity

and, therefore, an accurate level of expressiveness in the music performance the content of the vocal and facial expressions should, as discussed and in best cases observed, corroborate the emotional meaning of the musical piece, as it is revealed through the analysis of the musical structure of the song and of its poetic content. This was achieved in two different ways:

i) Through the performer self-observation and the 'eye-ball' analysis of the videotaped performances.

ii) Through the reading of the graph printed out from the computer, immediately after each performance, which showed the audience's real-time feedback of the perceived emotional meaning (see Figures XIX and XX).

The elaboration of the expressive performance tool was based on four complementary sections:

1. Performative-structural (musical-poetic) analysis: In order for the singer to support his expressive work, it seemed necessary that the musical work should be understood structurally. This section, necessary for the interpretation of Experiment-IV, can be found in Appendix F.

2. Computer software was designed to create a reliable way to register the audience feedback to the performer's interpretation. See Appendix G (software program text) + CD 4 (software for Experiments III and IV).

3. Real-time feedback from the audience of their perception of emotional sates in the singers performance: In order for the singer to learn how the audience perceives and understands the vocal and facial cues used, a graph was printed out immediately after each performance showing the audience's real-time feedback of the recognised emotional meaning (see Figures XIX and XX).

4. Videotaped feedback to the performer of his/her performance: In order for the singer to be able to learn how to control his/her expressiveness, videotaped feedback of the facial and vocal expressions of the singer allowed an 'eye-ball' analysis of the vocal and facial expressions used in the song performances, and also the possibility to compare the

videotaped feedback sequence of the performance with the audience feedback timings of the printed out graphs (CD 3 shows the videotaped performance-trials).

5.3.2 A Performance-Practice Application: a Case Study

Experiment-IV (see above, Chapter 4) constituted, as a case study, a practical application of this empirical tool. The ecological validity of this experiment (the audience was listening to an uncontrolled real piece of music rather than an excerpt or contrived auditory stimulus) together with the audience real-time feedback allowed the singer to check and improve the consistency of his performance's expressiveness, i.e., the accuracy of the different expressive cues used when performing the referred Schubert Lied. The presented audience feedback appeared in form of a graph showing to the performer, in a clear and simple way, how the audience received and comprehended the performer's music interpretation. The performance was simultaneously videotaped and the results and timings of the graphs could be compared and analysed through the videotaped performance and then reported to the musical structure of the song performed. The graphs presented time in seconds, and thus allowed a close check between the performed musical structure and the expressive means used by the singer. Facial and vocal expressions were reviewed, compared and analysed through timings and musical structure and then compared with the ensemble of the perceivers' answers according to their timings and emotional recognition. An immediately subsequent second performance enabled the singer to correct and improve the facial and vocal expressiveness, in order to communicate accurately and consistently the intended meaning of his interpretation of the musical structure. The second performance was again videotaped and new graphs were printed and analysed in order that the performer could be able to have a clear feedback of the improved vocal and facial cues and of the way his new performance was received and comprehended by the audience.

5.3.3 Successes and Limitations of the Techniques used in the 'Performer's Expressive Tool'

As can be understood from the simple observation of the computer's graphs printed out after the first and the second performance of the Lied 'Die Post' (see Figures XIX and XX), one of the major successes of the techniques used has been the clarification of the different intended emotional meanings expressed in the music performance and, therefore, a clear communication of the song's message and of its interpretation. These techniques

made not only the performer aware of the qualities, deficiencies and consistency of the expressive vocal and facial cues used in the performance, but also made him aware of the degree of accuracy of the audience's perception and recognition of the intended meaning communicated. In other words, the techniques used by the 'performance expressive tool' allowed the performer to correct and improve the expressiveness of his performance, according to two different parameters:

A) The accuracy of the feedback of the audience's answers to the expressiveness of the different performances, given by the performance graphs.

B) The consistency of the 'appropriate' vocal and facial expressions of the performer's interpretation with the selected structural features of the song during the performances presented by the videotaped feedback.

Ecological validity was another important success of the techniques used in this case study. The use of real music performance and of a real-time audience's feedback allowed for an uncontrolled concert situation and, also, to consider that the audience's feedback happened in a real-time perception reaction to the re-presentation of an emotional expression and not like a re-collection of the recognised emotional meaning.

Otherwise, the techniques used seem to present some limitations in terms of the continuity and variety of the emotions communicated during the performance, and the limited choice of emotions offered by the descriptive terms. It seems also impossible to consider any real-time information about the emotions felt during the performance. The only possibility to consider the side of the induced emotion communication, in this case, would be a qualitative interview to all the elements of the audience at the end of the performance. Surely, it is possible to suggest that the expression of a performance may arouse emotions in an audience through different kind of processes. One of these processes is certainly, and as an example, the emotional contagion, i.e., people get the emotions of others by hearing their vocal expressions or by watching their facial expressions (see Hatfield et al., 1994:5). According to Ekman (1993), there is enough evidence to claim that facial manipulation of expressions (for example, smiling, frowning) can evoke an associated feeling response.

However, the investigation undertaken here has been primarily concerned with how emotion represented in music performance is perceived and recognised by an audience.

So, in one way, the message of the song was clearly communicated through the use of the discreteness of the intended emotional meaning(s) analysed in the music structure, and through the consistency and accuracy of the selected and, later, improved performer expressive cues. In another way, a recognisable limitation of the study was the impossibility to register the continuity and variety of the more subtle emotional meanings of the piece. In fact, many of music's most subtle expressive qualities relate to structural changes across time, as it is the case of 'shifting emotions', 'blending emotions', and 'conflict emotions' (see Gabrielsson & Lindström, 2001). So, the main focus of the audience was compelled in this study to remain connected with the singer's strongest expressive moments and limited to the four categorical emotions strongly expressed. The nuances introduced during the song remained unanswered and, probably even, unnoticed. Equally, because the most relevant cues of emotional musical meaning remained focused on facial and vocal expressiveness, all the other hypothetical existent Non Verbal performance communication cues remained unassessed.

5.4 NOVELTY AND SUCCESSES OF THE STUDY

The portraying of an emotion in singing is, certainly, determined by a highly complex mixture of factors. Music, text, interpretation, context, voice technique and articulation, bodily movements, singer's personality and interactivity with the audience seem altogether, to play an important role in the music-emotional expressive outcome.

According to Manén (1974), historically, *Bel Canto* vocal technique was a musical exploration of the different vocal expressions for the different emotional states. So, in a practical way there has been an exploration of vocal emotion in the performance of singing of the Western Art Music Canon ever since opera was born in the 17th century, but few systematic empirical studies. The current study investigated empirically how voice, emotion and facial gesture might be connected in Western classical solo singing (within the Nineteenth Century German Lied repertoire). The need for such a study arose out of an awareness of several issues. Firstly, facial expression in singing is often discussed

anecdotally, but has rarely been subjected to any empirical analysis. Secondly, singing teachers often ask singers to 'sing with the eyes', 'make a smile' and so on, to achieve technical ends in singing. Thus, it was considered important to know whether these different facial movement expressions do in fact affect the quality of the produced vocalisation. Thirdly, given the second point, it seemed important to know if there was an objective correlation between the facial gesture and the sound made - in terms of its expressive intention. From a perceptual perspective, for instance, does the audience understand more if the singer looks as well as sounds 'sad', and what do these emotions objectively look and sound like? Fourthly, it is well known that singers and actors often show empathy with an emotional state, without entering into it completely. In fact, singers 'act' out emotions. It was a final intention of this work to explore the extent to which the emotional expressions requested were perceived as being authentic by both the performer him/herself and the audience. It was possible to match these data against 'norms' for emotional expression in the face by comparing the profiles of the singers with measurements of facial formation/musculature arrangement for real emotions recorded by Ekman & Friesen (1976) from still photographs taken when people were subjected to specific emotion eliciting situations.

Of the existing empirical work, key research has been undertaken by Kotlyar & Morozov (1976), Sundberg (1982), Sundberg et al. (1995), Gabrielsson (1995), Juslin (1997), Scherer (1995), Scherer & Siegwart (1995), Ohgushi & Hattori (1996), and Salgado (2000, 2001, 2002) who have demonstrated that when singers were asked to sing with different emotional intentions, say, of happiness, sadness, anger, fear, or even with no emotion at all, very different vocal sounds and vocal expressions emerge. For example, when asked to sing with the emotional intentions of happiness, sadness, anger, fear, and no emotion, very different spectrographic analyses of the vocal sounds emerge. In happiness, for instance, there is a much higher frequency than the other emotions, the tonal course of the pitches is moderate, both up and down, the tonal colour comprises many overtones and the volume is loud. In sadness, there is a much lower frequency produced, the tonal course of the pitches is downward, the tonal colour is very restricted, with few overtones and the volume is soft.

The existing amount of empirical work concerning the acoustic concomitants of emotional expression in singing is, though meagre, sufficient to pull out of it some conclusions about the way listeners infer emotion from vocal cues and how this set of cues is being used by performers in order to achieve communication of emotion. However, the complementary empirical research concerning the visual concomitants of emotional expression in singing, with the exception of Ohgushi & Hattori (1996) and Salgado (2000, 2001, 2002), is practically non-existent. These research projects have built up on the evidence on emotional facial expression (Ekman & Friesen, 1978), on the expression and nature of emotional facial expression (Ekman, 1984), on the argument and evidence about universals in facial expression (Ekman, 1989), about the universal and innate character of emotional facial expression (Izard, 1994), and the study and recognition of emotional expression using synthesised facial muscle movements (Scherer et al., 2000; Calder et al., 1996; and Bruce & Young, 1998) in everyday facial expressions. Given the almost complete absence of research precedents for what happens to the face in singing, it was hypothesised that the face would differ greatly according to the emotion, with happiness involving very different gestures to sadness, as it was felt that there would be a correlation between size of expression and loudness of sound produced. These hypotheses were in part based on intuitions from everyday observations, but also emerged out of drawing parallels with the work of Davidson (1994) who discovered that when a pianist was asked to perform with different emotional intentions, the louder he played, the larger his movements were in order to produce the sounds.

Linking these general research findings about musculature to singing technique, it is important to note that in classical singing, the intention is generally to keep the larynx free, to allow for optimum vibration. Additionally, the singer is taught to use the resonating cavities of the face and the pharynx. To achieve this, vowel sounds are often modified from those used in everyday speech, with the mouth opening rather more at the back than the front (Helmsley, 1998). These factors may have an impact on how the face works when the highly trained singer is asked to produce an emotional expression. In fact, there may even be some source of conflict, with natural facial expression involving a muscle in one direction which may need to function in another way for the sake of optimum vocalisation of the same emotion when interfaced with the technique of singing.

Basically, the research undertaken on the visual concomitants of emotional expression in singers' performance (Ohgushi & Hattori, 1996; Salgado, 2000, 2001) has focused only, and so far, on the qualitative analysis of the recognition of the emotion expressed in singing. This has been achieved by forcing listeners' responses to the perceived acoustical and visual elements of the emotion expressed through a set of adjective scale choice concerning the type of emotion that the listeners are supposed to recognise through them. And, by differential inquiring and rating scales about the intensity or mode (exaggerated, proper, deadpan or insufficient) by which the required emotion has been recognised. The only quantitative-objective known study focusing on the emotional facial expression in singing is made in this thesis. Thus, an important novelty of this thesis was to consider systematically the study of the facial expression of emotion in the performance of singing. Emotional facial expression in singing was investigated according to different perspectives and parameters:

a) It was investigated, as already stated, building up on, and by comparing with, the evidence on emotional facial expression (Ekman & Friesen, 1978), on the expression and nature of the emotional facial expression (Ekman, 1984), on the argument and evidence about universals in facial expression (Ekman, 1989), about the universal and innate character of emotional facial expression (Izard, 1994), and the studying and recognition of emotional expression using synthesised facial muscle movements (Scherer et al., 2000), in everyday expressions. (Experiment-III).

b) It was investigated by focusing on the qualitative-objective analysis of the recognition of the emotion expressed in singing, by forcing listeners to respond to the perceived acoustical and visual elements of the emotion expressed through a set of adjective scale choice concerning the type of emotion that the listeners were supposed to recognise through them. And, by differential inquiring and rating scales about the intensity or mode (exaggerated, proper, deadpan and less) by which the required emotion has been recognised. (Experiment-III).

c) It was investigated by focusing on the quantitative-objective results extracted from the emotional content of the facial expression of the different sung performances. From the results of this particular investigation, it seemed also possible to conclude that singing with different emotional meanings involves the production of facial expressions, which imply a set of very different and specific facial movements. These movements are directly connected with the intended emotional expression sung and can be detected and captured with the help of the suggested quantitative-objective procedure, a point-light technique used according to Bassilli's (1978) experimental work (and that Davidson, 1991, also used to investigate the perception of expressive movement in music performance). This procedure seems to certify not only of the existence of these movements but also of the specific differences of their behaviour (according to the different intentions of the singer) independently of the perceiver's subjectivism. (Experiment-I).

d) It has also been possible to link the quantitative results, through a qualitative analysis of the point-light displays, to a reliable emotional interpretation of the point-light displayed information, i.e., the information contained in the point-light display show similar characteristics to everyday emotional expressions. So, it seems possible to conclude that emotional facial expressions in music performance (at least the four basic emotions under study) do not only have clear distinct movements in themselves and between one another, as the information they convey, even in the subtlest display of their dynamic form, can be accurately perceived and recognised as categorical emotional meaning. It indicates that also in the artistic domain, the dynamic information is an important perceptual source for emotion, and it might be seen as an important complement to Bassili's (1978) investigation, because despite the relative perceptual subtlety involved in picking up information contained in his studies of the perception of emotion in the human face in everyday world, there has been a lack of information examining the artistic domain of emotional facial expressiveness, in which there are intentions of an expressive nature, which have to be contained within the constrained movements of technical and

physical adjustments – as in the case of musical performance. Therefore, despite the information contained in the point-light display being only visual, the study shows that the dynamic aspect of the visual information is an important part of the communication process. (Experiment-II).

e) These investigations used an ecological approach, which allowed for the verification and the assessment of facial emotional expression in the performance of singing under realistic conditions. It also allowed for an investigation of the degree of accuracy of the audience perception and recognition of facial emotional meaning communication in a real-time performance under a real concert situation and according to consistent expressive musical structure cue utilisation, i.e., it was the interaction between the analysis of the composed structure and the performance of the song that determined the emotional expression of the piece (see above the development of an expressive performance tool – successes and limitations of the techniques used). (Experiment-IV).

5.5   FINAL CONCLUSIONS, LIMITATIONS AND IDEAS FOR SUBSEQUENT STUDIES

This thesis investigated the four main universal and innate categorical expressions of emotional meaning. As has already been above referred, there are a number of limitations of this study, which might provide stimulus for further subsequent studies. One main limitation of this study was surely the limited number of emotions investigated. This limitation, however, was assumed from the start, as the necessity of the scientific validation of these four categorical emotional expressions within the artistic domain of the performance of singing. Of course, more subtle and diversified spectra of emotional nuances in music need to be investigated. Nevertheless, the work now available within the parameters of these four basic emotions may constitute the roots for a more specialised research. Amongst the more subtle emotional nuances of music remaining unexplored, should be referred those related to structural changes across time (shifting emotions, blending emotions, conflict emotions, etc) which require a much more detailed approach to the study of expressive performance cues. But, agreeing with Juslin (2001), it is possible to argue that, "ultimately, the real force of music's expressiveness – if not its uttermost subtlety – resides in such roots (the evolutionary framework of 'basic emotions')". Indeed,

the importance of these four basic emotions remains, as said, in the essence of their universality, innate and evolutionary character and categorical expressive behaviour. The discreteness of their content and the clearness of their perception and recognition, as well as the degree of the development and of consistency of the psychological studies and music investigations concerned with them, creates an important and consistent level of ecological reliability, which constitutes the necessary basis for any future scientific study.

Another limitation of this study connects with the equally important aspect of the musically induced emotions. In fact, the focus of this thesis has been concerned with the perception and recognition of the emotions represented through the facial and vocal expressions in singing. Surely, the expression of a performance may arouse emotions in an audience. It is the case in some of the already discussed processes, like the *emotional contagion* or the performer's underlining of some of the recognised structural features intrinsically connected with the generation of psychological phenomena like tension, tears, shivers, and thrills (Sloboda, 1998). The study of this communication of emotion in music performance requires different approaches and certainly different measuring instruments and techniques. Certainly a dimensional approach and the use, for instance, of EEG (Schmidt & Trainor, 2001) or physiological measurements techniques, such as heart beat, respiration, skin temperature, electrodermal responses (Krumhansl, 1997; and Nyklicek et al., 1997 for recent examples where physiological measures did discriminate among emotions) or, even, electromyography (Lundqvist et al., 2000). See Scherer & Zentner (2001) for a general review of physiological changes literature.

Surely, the use of only one or two singers and the limited number of perceivers in the empirical work undertaken practically reduce the practical investigation of this thesis to a case study. Also, the use in the investigation of only one style of music – the Nineteenth Century German Lied repertoire – reduces the range of the field of the research, especially because the most relevant cues of emotional musical meaning remained compulsorily focused on the facial and vocal expressiveness. So, because Lied is an art form of singing which is mainly concerned with the vocal and facial expressiveness of the singer, all the other hypothetical existent Non Verbal communication – bodily movements, posture, interpersonal distance (Lewis & Haviland-Jones, 1993) – and, therefore, all the other Non-

Verbal performance communication cues remain to be assessed. So, further and wider experimental work should be undertaken in order to confirm the empirical evidence now revealed. But, it must also be added that, from the point of view of the investigation it has been also assumed that it was primarily important to reduce the field of investigation in order to focus the practical experiments and thus the related conclusions in a workable amount of data to be analysed and interpreted.

So, as an aim and resulting ideas for subsequent studies, it should first be considered to extend the study of emotional expression to other and more subtle emotional categories and meanings. Secondly, the study of emotional meaning should be also developed in the sense of connecting the expressive elements of music performance with the structural elements of musical analysis, also where more subtle forms of emotional meaning are communicated, a research to which Juslin has made a precious contribution (see Juslin, 2001). Thirdly, the combined study of the dimensional and the categorical approaches to emotional meaning in music performance might be also highly productive if it could find a way to incorporate in the same investigation the contribution of the study of the physiological reactions. Scherer & Zentner (2001:374) refer exactly this possibility by claiming that the issue becomes then "to demonstrate which of the basic emotions are elicited by a specific piece of music, as judged from the physiological response pattern observed". Thus, in other words, it might be possible to include the empirical evidence revealed by the study of the arousal concomitants of emotion within the study of the perceived and recognised categorical emotional contents of the emotional meaning communication in music performance. Krumhansl (1997), for instance, showed statistically significant differences between three music emotional meanings: sadness, fear, and happiness. Sadness excerpts produced large changes in heart rate, blood pressure, skin conductance, and temperature; happy excerpts produced large changes in respiratory patterns; and fear excerpts, produced large changes in pulse transit time and amplitude. Bartlett (1996) showed that, as predicted by Ekman (1972), there was increased *corrugator* activity (frowning) while listening to music with negative valence and increased *zygomatic* activity (smiling) for music with positive valence. These results are consistent with Lundqvist et al.'s (2000) study, in which autonomic activity and facial electromyography were measured while an audience was listening to sad and happy music. Thus, the

interaction of a multi-modal approach where the study of the perception and recognition of categorical emotional meaning could be allied to physiological discriminative emotional measurement techniques could contribute with important information in what concerns the inducing of emotions through vocal and facial categorical expressiveness.

An idea for a subsequent study would be, for instance, to consider a way of monitoring the audience's facial expressions to see if they mirror those of the singer. This could be particularly useful to understand and differentiate at which level the emotional response to music has been identified by the audience. It also could aid to clarify some possible confusion between the perception and production of emotion. As previously related, in Experiment-IV, and especially connected with the communication of the emotional meaning, with anger, it seemed that a sort of confusion appeared, even if in a small way, during the audience response to 'Second run through' (see Figures XX and XXII). In fact, some of the responses of the audience seem to have connected the last performed emotional peak of the song to fear and not to anger, as was supposed to happen according to the singer's interpretation; this was probably because they confused the emotion perceived with the emotion felt. Indeed, Zentner et al. (2000), in a study focusing on the difference between 'felt vs. perceived' music emotional contents, noticed that the type of instruction given to the audience influenced the type of response given and changed significantly the results. They conclude that the distinction between perception and induction of emotion is an empirically valid one, and surely it will be essential to find empirical approaches that can aid to take this distinction into account.

Despite the great amount of information gathered in the field of the emotional communication in music performance, it might be said that the study of the relationship between music and emotion is still at a preliminary stage. The current thesis shows that a multi-modal approach (in the thesis case, a categorical approach with complementary Verbal and Non Verbal reports, ecological valid approaches, and expressive behaviour quantitative measurements) is the most productive way forward. Linking these aspects with all the different aspects of emotional investigations in music performance referred (continuous response approach, dimensional approach, open-ended or free verbal approaches, study of the induced emotion, emotion physiological studies, etc.) will

certainly aid to expand the information collected in this thesis, and hopefully bring to the research of the expressive movement in music performance the necessary information that will assist the performer to comprehend fully and improve the expressive elements of his/her performance practice. The empirical work of this study also suggested its relevance by the direct implication in some practical applications for singing and the Teaching of singing, as the elaboration of a 'performer's expressive tool', which could bring some aid to the singer's emotional expressiveness skills. Also, because singing connects emotion facial expression with music emotional character, the study here undertaken suggests the possibility of another practical application – a 'therapeutic emotional tool' - for the research with brain damage patients that have lost the capacity to recognise certain emotion expressions, like fear or anger. Neuropsychological research could find in the emotional expression in the performance of singing a strong vehicle for emotional representation, once music as an aural stimulus could be assisting the visual stimulation generated by the facial expressed emotion.

This thesis has primarily been an attempt to contribute and develop empirical methods to the investigation of the communication of emotional meaning in music performance (singing). By investigating the relevance and interface of the visual and acoustical dimensions of the expression of emotion in music performance there has also been an attempt to contribute to the visual component of expression in music live performance, a field of investigation recently opened up by Davidson (1991, 1993). In this sense, it appears to be a pioneering investigation by exploring systematically the emotional meaning expressed through the face and the voice of the singers within the performance of singing. Once the aim of a performing musician is to communicate expressively with the audience, the accurate audience perception of the intended emotional meaning expressed is implied. This thesis hopes to contribute to a better comprehension and knowledge of the expressive means that constitute the performer's (singer's) expressive cues, and to a deeper understanding of the relationship between music and emotion.

APPENDICES

Appendix A: The Emotional Communication Test (Experiment II)

Encircle the emotion category which best fits the stimulus:

1)   Neutral Sad Happy Angry Fearful Surprised Disgusted

2)   Surprised Disgusted Neutral Sad Happy Fearful Angry

3)   Disgusted Sad Angry Surprised Fearful Neutral Happy

4)   Fearful Happy Surprised Sad Disgusted Angry Neutral

5)   Sad Surprised Happy Disgusted Angry Neutral Fearful

6)   Happy Disgusted Angry Neutral Surprised Fearful Sad

7)   Angry Sad Disgusted Fearful Happy Neutral Surprised

8)   Surprised Disgusted Neutral Sad Happy Fearful Angry

9)   Sad Surprised Happy Disgusted Angry Neutral Fearful

10)   Disgusted Sad Angry Surprised Fearful Neutral Happy

11)   Neutral Sad Happy Angry Fearful Surprised Disgusted

12)   Angry Sad Disgusted Fearful Happy Neutral Surprised

13)   Fearful Happy Surprised Sad Disgusted Angry Neutral

14)   Happy Disgusted Angry Neutral Surprised Fearful Sad

15)   Neutral Sad Happy Angry Fearful Surprised Disgusted

16)   Disgusted Sad Angry Surprised Fearful Neutral Happy

17)   Angry Sad Disgusted Fearful Happy Neutral Surprised

18)   Surprised Disgusted Neutral Sad Happy Fearful Angry

19)   Happy Disgusted Angry Neutral Surprised Fearful Sad

20)   Sad Surprised Happy Disgusted Angry Neutral Fearful

Appendix B: The Emotional Communication Test (Experiment III)

Encircle the category which best fits the stimulus perceived in the videotaped excerpt:

Neutral   Sad   Happy   Angry   Fearful   Surprised   Disgusted

and

Rate each stimulus by encircling the number which best fits the level of performance expression: (1 deadpan; 4=normal ; 7= exaggerated)

deadpan			normal			exaggerated	
1	2	3	4	5	6	7	

SS1

A   Neutral Sad Happy Angry Fearful Surprised Disgusted

2   3   4   5   6   7

B   Surprised Disgusted Neutral Sad Happy Fearful Angry

1   2   3   4   5   6   7

C   Disgusted Sad Angry Surprised Fearful Neutral Happy

1   2   3   4   5   6   7

SS2

A   Fearful Happy Surprised Sad Disgusted Angry Neutral

1   2   3   4   5   6   7

B   Sad Surprised Happy Disgusted Angry Neutral Fearful

1   2   3   4   5   6   7

C   Happy Disgusted Angry Neutral Surprised Fearful Sad

1   2   3   4   5   6   7

SS3

A   Angry Sad Disgusted Fearful Happy Neutral Surprised

      1   2   3   4   5   6   7

B   Surprised Disgusted Neutral Sad Happy Fearful Angry

      1   2   3   4   5   6   7

C   Sad Surprised Happy Disgusted Angry Neutral Fearful

      1   2   3   4   5   6   7

SS4

A   Disgusted Sad Angry Surprised Fearful Neutral Happy

      1   2   3   4   5   6   7

B   Neutral Sad Happy Angry Fearful Surprised Disgusted

      1   2   3   4   5   6   7

C   Angry Sad Disgusted Fearful Happy Neutral Surprised

      1   2   3   4   5   6   7

SS5

A   Disgusted Sad Angry Surprised Fearful Neutral Happy

      1   2   3   4   5   6   7

B   Neutral Sad Happy Angry Fearful Surprised Disgusted

      1   2   3   4   5   6   7

C   Angry Sad Disgusted Fearful Happy Neutral Surprised

      1   2   3   4   5   6   7

Appendix C: The Emotional Communication Test (Experiment III)

NON VERBAL RECOGNITION MODE (in all three modalities: acoustically and/or verbally)

Encircle the facial expression which best fits the stimulus perceived in the videotaped excerpt:

and

Rate each stimulus by encircling the number which best fits the level of performance expression:

(1= deadpan; 4=normal ; 7= exaggerated)

deadpan			normal			exaggerated
1	2	3	4	5	6	7

SS1

A   Neutral Sad Happy Angry Fearful Surprised Disgusted

1   2   3   4   5   6   7

B   Surprised Disgusted Neutral Sad Happy Fearful Angry

1   2   3   4   5   6   7

C   Disgusted Sad Angry Surprised Fearful Neutral Happy

1   2   3   4   5   6   7

SS2

A   Fearful Happy Surprised Sad Disgusted Angry Neutral

1   2   3   4   5   6   7

B   Sad Surprised Happy Disgusted Angry Neutral Fearful

1   2   3   4   5   6   7

C   Happy Disgusted Angry Neutral Surprised Fearful Sad

1   2   3   4   5   6   7

SS3

A   Angry Sad Disgusted Fearful Happy Neutral Surprised

         1   2   3   4   5   6   7

B   Surprised Disgusted Neutral Sad Happy Fearful Angry

         1   2   3   4   5   6   7

C   Sad Surprised Happy Disgusted Angry Neutral Fearful

         1   2   3   4   5   6   7

SS4

A   Disgusted Sad Angry Surprised Fearful Neutral Happy

         1   2   3   4   5   6   7

B   Neutral Sad Happy Angry Fearful Surprised Disgusted

         1   2   3   4   5   6   7

C   Angry Sad Disgusted Fearful Happy Neutral Surprised

         1   2   3   4   5   6   7

SS5

A   Disgusted Sad Angry Surprised Fearful Neutral Happy

         1   2   3   4   5   6   7

B   Neutral Sad Happy Angry Fearful Surprised Disgusted

         1   2   3   4   5   6   7

C   Angry Sad Disgusted Fearful Happy Neutral Surprised

         1   2   3   4   5   6   7

Appendix D: The Audience Questionnaire-Interview

Please complete:

1 - Your age: ....
2 - Gender: ....
3 - Occupation: ....

4 - Have you received musical training?

5 - Have you heard this song before?

6 - . How did you feel about this interpretation?

    Prompt: consider the following:

    - Did the emotion conveyed by the singer fit the emotional state of the character of the song.

    - How authentic were the emotions conveyed by the singer?

7 - What strategies were you using to make your decisions?

8 - How did you find the test? Prompt: easy/difficult/neutral to undertake?

Appendix E: The Performer's Questionnaire - Interview

1 How did you approach this task?

2 What were your thoughts and feelings at that time?(where you convinced of the accuracy of your different levels of performance: deadpan, just, exaggerated?)

3 What are your thoughts and feelings now as you view the video? (Surprises? Disappointments? Is the normal performance as you expected it to be? Is the deadpan/exaggerated performance approximately under/over stated?)

4 Did you feel the emotions you intended to perform?

5 Do you believe to have conveyed the emotions of the character of the song?

6 Can you say what you might do differently to clarify or modify your interpretation in some way?

7 How might you think about doing these particular interpretations differently now?

8 In what ways do you use your technique and your real understanding of emotion to develop your performance?

9 How focused were you on your performance?

10 In what ways do you think or feel you could use this video approach and the results of the perceiver's inquiry to help you to develop your communication and interaction of the emotion within the art of singing?

11 Where you aware of the facial muscle activity you had to produce in order to get the required emotion expressed?

12 Think of the four fundamental emotions you have been dealing within this experiment (sadness, happiness, anger or fear): Would you be able to perform vocally, within singing, any of the required emotions expressing facially another one?

13 Do you feel that the degree of accuracy on the communication of the emotion is dependent on the authenticity of the emotion expressed?

14 Within the performance, to what extent did you actually feel the following emotions: sadness, happiness, anger, and fear.

Summary of responses: Both singers believed that their emotions were 'authentic' in that they were constructed out of their memories of those states. However, they were all able to recognise that in their different interpretations some were more 'successful' performed than others.

Appendix F: Performative Structural Analysis (Experiment IV)

Appendix F is constituted of three interrelated items:

A. A small introduction, in order to contextualise the Lied *Die Post* within the 'Winterreise' cycle's atmosphere and within the general Romantic content of Schubert's Lied work.

B. An interpretative-structural analysis of Schubert's '*Die Post*', aiming to assist and examine in further detail the investigation undertaken in Experiment IV.

C. Analysis of the vocal and facial expressive cues to be used by the singer during the performance, aiming for a clear and accurate communication of the emotional meaning of the song.

D. Score of the Schubert song: "*Die Post*".

A. Shubert's 'Winter Journey'

Winterreise op. 89 (1827) by Franz Schubert

The texts from *Die Winterreise* by Wilhelm Müller which Franz Schubert set to music in 1827, a year before the poet's death, mirror the Wanderer's myth whose demons take us to the deepest and coldest regions of the human condition. Schubert would have presented these Lieder to his friends at Schober's house as a cycle of horrifying songs. Influenced by Greek and Roman literature, folklore, opera, and German drama, and British and German poetry of XIX century, Müller wrote this cycle of poems between 1822 and 1824. *Die Winterreise* is characterised by Romantic imagery and by the archetypes like the myth of the 'Wanderer' and of 'Winter travel', favourite topics of the Romantics, and draw the story from the most ancient civilisations. The language used by the poet, his appreciation of music theatre and the visual arts, as well as the British and German Romantic tradition of XIX century, offered up fertile ground for the creation of *Die Winterreise*. From the several influences he had, Blake's mythical poetry was probably the most powerful. Here, Müller assimilated the idea of a divided human *psyche*, the spectre of a journey through life trying to meet its shadow. Müller conceived his poems to be sung, as he states in his diary in 1815: "I can't either play or sing, however, when I write verses I do sing and play. Producing melodies my songs can be more agreeable. But courage! Maybe there is a bright spirit somewhere that could listen to the tonalities of the words and bring them back". This bright spirit was Schubert who wrote the music for the author's *Die Winterreise* as well as for *Die schöne Müllerin*.

The composer found Müller's poems in Franz Schoeber's bookshop in the winter of 1826. Müller had published his poems by stages, the first twelve in Urania: *Taschenbuch auf das Jahr* 1823, then in 1823 in the *Deutsch Blätter für Poesie und Litteratur, Kunst, und Theatre*. In 1823 he published the whole cycle comprising *Die Post* and *Täuschung* in the *Waldhornisten II*. A few years later, in 1827, the composer came across some of these poems, and so born the first twelve Lieder of *Die Winterreise*. A few months later Schubert discovered in the second book of W. Müller's *Poems taken from papers of an itinerant horn player* his own "Sequel to the Winterreise", with another twelve poems. So, all together twenty-four songs, the cycle was complete. Schubert started working on this cycle

of songs without knowing all of Müller's poems. This fact and by using Müller's latest drafts, made him break with the musical structures he worked out in the meantime. Often seen as autobiographical, this cycle should also be metaphorically understood and explored as a strong psychological articulation of the romantic philosophy and myths. The *Winterreise* is a dream, a vision of travel that the *wanderer* makes through the mind and heart to the depth of the soul. As action takes place in the human *psyche*, the wanderer surpasses himself drawing on elements from nature as symbols of unquestionable truths of the human soul. Through the *Winterreise*, as well as all Romantic literature, nature mirrors man. The musical and metaphysic progressions turn out to be manifestations of psychological travels. In fact, in this cycle of songs, the depth and nature of the feelings and emotions expressed are intensely experienced by the composer, being musically translated with a unique mastery, revealing a strong musical technique and rationality In this work, disillusion always anticipates action with the relationship between the different states of soul, emotions and feelings and the different elements integrating the narrative. Schubert translates several physical and psychological emotions mainly in the interaction between the piano and the voice. Many and diversified are the technical and stylistic devices used in *Die Winterreise*: the melodic control, the nature, subtlety and singularity of the modulations; the harmonic mastery; the way melody and harmony are used and combined, and expressively intoned; the constant change of mode and tonality; the preponderance of dactyl rhythm, known as the rhythm of death: and the repetition of folkloric melodies are just some of the multiple examples used by Schubert to translate the Wanderer's spirit. In Winterreise the simplicity of the music enhances the nature of the feelings. Schubert uses rather simple and concise musical and thematic structures; the evolution and discursive development is grounded in the variation; the progressive change of vocal elements from singing to declaiming are resources used in the expressive intensification of each work. Schubert's music, describing and faithfully translating the myth and the emotions and the feelings of the traveller gives a peculiar dramatic strength to the cycle, making of it an example of perfection in the way the composer visualises them and changes them into sonorous realities. The relationship between the different states, emotions and feelings of the traveller, and the different elements that make part of the narrative is noteworthy.

The cycle *Die Winterreise* is, in fact, a true journey to the inner and dark side of the human soul, coloured with the dramatic shades of loneliness, despair, and the restlessness of death itself, a true trip to the no-man's land of the human soul. The 24 songs that constitute this cycle are:

Gute Nacht (Good Night),

Die Wetterfahne (The Weather cock),

Gefror'ne Tränen (Frozen tears);

Erstarrung (Chill torpor),

Der Lindembaum: (The Linden tree),

Wasserflut (Floodwater),

Auf dem Flusse (On the river);

Rückblick (Looking back),

Irrlicht (Will-o'-the-wisp),

Rast (Rest),

Frühlingstraum (Dream of spring),

Einsamkeit (Loneliness),

Die Post (The Post),

Der Greise Kopf (The Grey head),

Die Krähe (The Crow),

Letzte Hoffnung (Last hope),

Im Dorfe (In the village),

Stürmische Morgen (The Stormy morning),

Täuschung (Illusion),

Der Wegweiser (The Sign post),

Das Wirthaus (The Inn),

Mut (Courage),

Die Nebensonnen (The Mock suns),

Der Leiermann (The Hurdy-gurdy man).

The aim of Experiment-IV was, as said, to investigate if the performer, according to the analysed structural meanings, was able to communicate accurately, during the continuous development of the musical structure of the song, all the previously selected emotional contents. In order that the interpretation of the song – *Die Post* – could be performed according to the music structural meanings, an interpretative analysis of this song is presented below, describing some of the Schubert's compositional devices. Figure XXIII, showing the interface of the three elements – music structural features, performance facial expression peaks, and accuracy of perceiver's recognition of the intended emotional meaning performed – can be seen in Chapter 4, Experiment IV, pg. 198. [49]

B. AN INTERPRETATIVE-STRUCTURAL ANALYSIS OF SCHUBERT'S 'DIE POST'

The full score of the song "*Die Post*" can be found at the end of Appendix F. It was decided to present the score of the song *Die Post* in the original version – E flat Major, and the structural analysis has considered this tonality as the main reference. However, because Toni is a Bass-Baritone, the song was performed in B Major.

Background of the song

'Die Post' is from Schubert's 'Winterreise', and in the first version of the cycle, this song was inserted after *Der Lindenbaum* (The Lime Tree). In the final version *Der Linderbaum* is the fifth poem of the first part, *Die Post* being the beginning of the second part of the cycle. By creating a dialogue between the voice and the instrument, between the traveller and the messenger's horn, Schubert changes the sound climate and atmosphere of the cycle making of this Lied the bearer of an expectation, potential hope and happiness. Here, the traveller gives outward expression to his thought, waiting for his lover's response. Turning to outside concerns, he looks forward to having a new strength for his life. The loneliness, which is overwhelming at the end of the first part of this cycle and shown in *Einsamkeit* (n°12), precipitates a turning point, which takes place in *Die Post* (n°13). More joyful and cheerful, this song impels the singer to create another colour and dynamic.

---

[49] Of course, it is impossible in the scope of this appendix to analyse all the 24 Lieder, as it will be done for the song *Die Post*, in order to exemplify all the creative devices that Schubert used in the composition of the cycle *Die Winterreise*. For a detailed information on Schubert's technical and stylistic devices used in this cycle see Feil (1975), Dommel-Diény (1976), Beaufils (1982), Kramer (1984), Bauer (1915), Capell (1957), Fischer-Dieskau (1971).

In E flat major, presents two brief cadences in D flat major and G flat major. The change of mode in the second and fourth stanzas creates new shades and discursive dynamics, leading to a contrast of colours and shadows, the hope and disillusion of the traveller, who deprived of his lover's presence looks forward to meeting her again. In terms of the music, Schubert makes use of musical and thematic structures in a discursive development mostly based on variation. Through the progressive transformation of the vocal elements, from the melodic to the declamatory, an expressive intensification is created. The melodic control, the subtlety and singularity of the used modulations, and the way melody and harmony combine, make this work unique.

General Overview of Structure

'Die Post' follows a strophic pattern, which can be interpreted as indicating the obsessive psychological state of the character of the song: "Immer wieder muss ich zu dir kehren und warten…, und hoffen…, obwohl (ich weiss es) von dir bekomme ich nur dein Schweigen!" (Again, and again, must I turn to thee, and wait…, and hope…, even knowing that from you I'll get nothing but Silence!) Therefore, the variations, though very small, happen due to the prosody. In melodic terms, each sentence plays a different role to emphasise the elements of the text. The song *Die Post* has the strophic pattern (A/B + A'/B'). Having not found any substantial difference that would require a new emotional interpretation, between the two first stanza (A/B) and the two last stanza (A'/B') of the song, musically or poetically, the analysis presented here deals only with the emotional interpretation of the first two stanza, i.e., the (A/B) part of the song. However, given the already spoken obsessive character of the song (through the use of the strophic pattern) a reinforcement of the emotional meaning expressed in A/B could be expected within the two last stanza (A'/B'), through the intensification of the same vocal and facial cues used for the first two stanza of the song.

Analysis of the structural-expressive features of the song

Bar 1 to bar 12

13.
Die Post.

The postillion's call announces the arrival of the mail coach. The traveller's heart has a moment of hope and excitement for the possibility of receiving news from his beloved. At the beginning of the song, Schubert announces the arrival of the post through a light and bright musical texture. The rhythm 6/8 and the hint Etwas geschwind (a bit lively), gives this song a gentle excitement peculiar to one who is waiting for his lover's news. The first sentence from the first stanza, Von der Strasse her ein Posthorn klingt (I'm listening the postman's horn in the street), is articulated on two pitches (B flat 3 and G 3), making an interval of 3rd major (the announcement of the postman's horn), according to what might be called the "descriptive-mimetic" character of the musical structure (see Iconic-structural-or-dynamic-extroversive coding in Chapter 3, pg. 111 and 123). It should be noted that the beginning of the work is based on the pianist's articulation of the notes in the I chord of E flat major, the tonic key, which develops through an extension of nine bars until it meets the singer's vocal line. The consonant character of the notes in the I chord of E flat major and the staccato articulation of these notes on the piano correspond

structurally to the atmosphere of expectation, gaiety, and agitation. The perfect cadence [V → I] that resolves in bar 9 (with the beginning of the vocal part) appears already anticipated in bars 7 and 8. An alternation in tonic/dominant E major chords at the words *Von der Strasse her ein Posthorn klingt* underlines the excitement of the character of the song, announcing the arrival of the postman.

From bar 12 to bar 16

The second sentence (bars 12 to 16), *Was hat es, dass es so hoch aufspringt, mein Herz?* (What is happening for you to beat so much my heart?), develops in a predominantly ascending way on the main notes of the harmony (V chord of E flat major, the tonic key), underlining the 'jumping' and 'bursting out' character of the text 'hoch aufspringen'. An alternation between tonic/dominant seventh chords at the sentence - *Was hat es so hoch aufspringt,* brings the musical structure to D flat major at bar 15. Apart from the *anacrusis* (on an interval of 2nd minor) the melody develops between the disjoined intervals of a perfect 4th, a major 3rd, a minor 3rd and a major 6th, evoking and re-confirming the 'emotional' character of the musical structure at this point (the uneasiness of the traveller's heart and the hope and joy of having news from his lover), here reinforced through other elements of the musical structure: tempo (etwas geschwind), dotted rhythm (varied), and articulation (staccato) – (see, for instance, for 'tempo' Scherer & Oshinsky, 1977; Juslin, 1997; Thompson & Robitaille, 1992; Krumhansl, 1997; Balkwill & Thompson, 1999; for 'rhythm' Wedin, 1972; Thompson & Robitaille, 1992; and for 'articulation', Hevner, 1937; Rigg, 1939; Juslin, 1997). Another element, consistent with the other structural features, is the use of the major sixth (A flat 3 → F 4), at bars 14/15, which is a structurally salient moment, indicating happiness as in many other cases in the music literature (see Cooke, 1959: 65). All these expressive elements of the musical structure may be considered to be a part of the *Iconic-structural extroversive coding* as referred in Chapter-3. The fact that the

long note on the word *Herz* (bars 15/16) corresponds to a point of maximal tension (f >), and points to a reasonable progress towards a goal (arriving at the tonality of D flat major), seems also to corroborate this idea of a pleasurable feeling of 'attending a goal' (for instance, receiving a lover's letter). Thus, in bars 9 to 16 Schubert stabilises the upper part of the musical texture, creating instead some instability in the bass (alternation of pitches at almost every new beat), and thus transferring to these new sounds the traveller's uneasiness and excitement *Was hat es, dass es so hoch aufspring, mein Herz?* (What is happening for you to beat so much my heart?). All these elements characterise and reveal the traveller's anxiety, restlessness and uneasiness, born out of the expectation and joy of getting news from the belovèd (of being loved).

From bar 17 to bar 25

The third musical sentence, on this very same text (bars 17 to 25), has a mainly descending characteristic on the principal notes of the harmony (6/4 position of the I chord of E flat major, indicating a lack of stability) and ends up in a written appoggiatura (bar 20). This fact (descending melodic direction and pitch contour – see Gerardi & Gerken, 1995 and Scherer & Oshinsky, 1977) announces musically the misfortune (the lack of news from his belovèd), and the inconstancy of feelings, which vary from joy to sadness, sureness to uncertainty, and which lead him to a distressing and fearful situation. A threatened self-preservation goal is clearly perceptible, a key appraisal for fear (Oatley, 1992), as the*Wanderer* is facing the possibility of not receiving news from his beloved. Musically

speaking, the tension, the anxiety and the fear are emphasised by chromaticism (see Krumhansl, 1997) and by the use of very small intervals, which always presuppose a dissonant harmony, reflecting the traveller's emotional mood. In a rhythmically contained texture, Schubert creates (bars 16 to 23) two opposite melodic movements: an ascending movement of major third and another descending of minor third. Together with a pedal (octave) from the bass (bars 19 to 23), this holds back the musical discourse and brings it suddenly to a rather quiet but dramatic focus. The obsessive pedal in the bass (B flat 2) throughout the five bars suggests the threatening lover's refusal, and can still be heard, as an echo of the traveller's frightened heart beat, in the second beat of bars 24 and 25. The melody suggests disillusion and hope (the goal conflict) in the type of intervals that are being used. For instance, in the two exclamations *mein Herz, mein Herz!* (my heart, my heart!), bars 21 to 25, Schubert creates two distinctive melodic movements within the character of a minor 3rd and carrying out a symmetry concerning the intervals used:

$$2^{nd} \text{ minor} \rightarrow 2^{nd} \text{ major} \leftarrow | \rightarrow 2^{nd} \text{ major} \rightarrow 2^{nd} \text{ minor}$$

This effect creates the goal conflict characteristic of fear (Oatley, 1992): will (or not) the postman deliver the news from his belovèd? Does she love him (or not)? He wants to know, but he is afraid to ask... (He wants to ask, but he is afraid to know the answer...) Musically, this is suggested through the attraction to the first degree step, at bars 23/24 (D natural 4 → E flat 4) representing the wish of receiving the letter (being loved), and the fear of no letter being delivered (the lover's refusal), symbolically designed by the upward melodic movement at bars 23 and 24 (B flat 3 → C 4 → D 4 → E flat 4) – see Scherer, 1977 - and through the appoggiatura at bar 20 (C 4 → B flat 3) and at bar 21 (D flat 4→ C natural 4), a form of accentual anticipation usually associated with racing heart (Sloboda, 1998), representing here the tension, and the fear. Thus, the second exclamation of *mein Herz* (bars 23 and 24) emphasises even more the emotional character of the frightening fact that the traveller has to face. The expectation created by this sentence is also expressed in the accompaniment through the musical dynamics and structure. Dynamically, the *crescendo/diminuendo* (<>), followed by the two subsequent (*fp*) effects, accentuates and emphasises the tension, drama, and fear (Scherer & Oshinsky, 1977) of the moment. The chromaticism, the dissonant harmony and the density of the musical texture used are

recognised musical factors for unpleasantness (Wedin, 1972), tension (Nielzén & Cesarec, 1982) and fear (Krumhansl, 1997).

Bar 26 – One bar of silence.

The excitement present at the beginning of this piece dissolves at the end of the first stanza where in a bar of silence (bar 26) an enormous dramatic tension builds. Schubert exemplifies here masterly the peculiar effect of silence in music according to Scruton (1997):

> "We hear silence as 'Schweigen', a being-silent. It is not cessation of action, but action of another kind - refraining, withholding, refusing."
>
> (Scruton, 1997:333)

From bar 27 to bar 36

Breaking up, the initial rhythm (bar 27) induces the melancholic and introspective character of the next stanza, which is reinforced with the movement to the minor mode and structurally confirms and emphasises the emotional transition to sadness (Hevner, 1936;

Rigg, 1937; Kleinen, 1968; Wedin, 1972; Crowder, 1985; Krumhansl, 1997; Peretz,
Gagnon & Bouchard, 1998). Schubert imposes a depressed atmosphere with a
preponderance of the rhythm of death. The deep sadness of the lover is revealed in the
containment and economy of musical means used from this moment on. The rhythm, the
harmony and the melody from an almost absolute constancy concur to emphasise the
emotional and psychological tension. The simplicity and subtlety of the musical writing
and of the resources used enhances and emphasises the protagonist's feelings and
emotions. The incessant 6/8 rhythm coupled with the simplest modulations gives the
listener the sensation of a 'lachrymose' feeling. The first sentence, *Die Post brings keinen
Brief für dich* (the postman does not bring any letter for you...) is static and develops
inside the interval of 2nd major (G flat 3 → A flat 3). A drop to D flat major creates a more
distant feeling, insinuating perhaps the typical and strange Romantic 'pleasurable' feeling
of longing for an inaccessible love. The second sentence, *Was drängst du denn so
wunderlich, mein Herz, mein Herz?* (Why are you so strangely restless, my heart?),
develops equally in a restrained way, arriving with the end of *wunderlich* at G flat major,
which temporarily becomes the harmonic reference, insinuating musically even more the
strangeness of this feeling for an inaccessible love. The melodic flow stabilises leading to
two pitches which create the interval of major 2nd ascending (A flat 3 → B flat 3), bars 30
and 33. The musical discourse creates a melodic progression of a major 3rd, from G flat 3
(bar 28) to A flat 3 (bar 30) and B flat 3 (bar 33) – characteristic of a musical factor
associated with tears (Sloboda, 1998) - and the constant return to a supporting note (like a
lament), characteristic of all the melodic movements in this second stanza. This deep
feeling of sorrow and despair culminates at bars 34 and 35, where, once more, the words
mein Herz, mein Herz (my heart, my heart) appear within a melodic line that creates the
intervals of a descending minor 3rd (E flat 4 → C 4) and an ascending minor 2nd (C 4 → D
flat 4) underlining the melancholic feeling of the traveller (see Cooke, 59; Maher &
Berline, 1982), working musically as a sigh. Immediately after (bar 36), the piano solo has
the benefit of the sharpened third (G flat 4), heightening the drama by echoing the
traveller's melancholic strangeness (the echo of the traveller's sigh, bars 36/37)...

From bar 37 to bar 46

The arrival of the dominant 7th of E flat major (bars 37/38), alternating with tonic (6/4's) and the double dominant of B major, brings the resolution to the tonic (E flat major) and again the word *mein Herz* appears, but this time bringing an outburst of the traveller's anger (bars 44, 45 and 46). By facing his frustrated plans and the demeaning offence against the tenderness of his heart, the traveller's feelings are now of disillusion, and his reaction of anger. The emotional interpretation – anger - at this point of the song is supported by Imberty's (1979:12) research, who suggests that the ascending melodic lines occurring in the alternation of dominant modulations are most times connected with the expression of violence or strong emotions (Iconic-structural/dynamic intoversive coding). It should also be noticed that the various culminating points of this song take place in different moments of the even and odd stanzas. The idea of anger is also supported by Nielsen's (1983) research, confirming that when the harmonic, melodic and rhythmic complexity is high it indicates the increased presence of tension and aggressiveness. Loudness is another musical factor that corroborates the idea of anger (see Juslin, 1997, 2000), and this point of the musical structure (bars 44/45 – A/B) and its later repetition in bars 89/90 (the A'/B' part of the song) are, indeed, the loudest moments of the song.

In the odd stanzas the highest culminating point (F 4) takes place in the exclamation of happiness *mein Herz* (bars 15, A part of the song) and (60, A'). In the even stanzas, the exclamation of anger is the highest point (F 4 → A flat 4), postponed to the next-to-the-last *mein Herz* (bars 44/45, B part of the song) and (89/90, B'), which also are the highest dynamic points of the musical structure (f >). This reveals that the timbre and the colour, thus, the character of the musical sound can be modified not only through the changes of mode, major/minor, but also through the registers used.

C. ANALYSIS OF THE VOCAL AND FACIAL EXPRESSIVE CUES (TO BE) USED DURING THE PERFORMANCE

Considering the different elements of the musical structure analysed, four main structural moments appear to be, in this song, directly connected with four possible emotional meanings, which are in appearance order: happiness, fear, sadness, and anger. The existing empirical evidence on music/performance and emotion, and the results obtained in Experiments I, II and III seem to suggest that the singer should try in the performance of the song the following expressive cues, according to the different emotional contents to be communicated:

Interpretation – Happiness

According to the performer's interpretation of Schubert's instruction at the beginning of the song *Etwas geschwind*, this moment occurred around the 20th second of the performance. The vocal and facial expressiveness developed from the beginning of the song to an expressive goal, which appeared clearly and strongly with the first emotional peak at the beginning of bar 15. This emotional peak corresponded to Toni's attempt to express vocally and facially the emotion of happiness, expressed in the words *mein Herz*, and within the 6th major jump (A flat 3 → F 4).

Vocally: In this condition, the vocal cues involved should include according to the empirical evidence of different investigations: High sound level (Juslin, 2000); bright timbre (Gabrielsson & Juslin, 1996); fast mean tempo (Gabrielsson & Lindström, 1995); fast tone attacks (Kotlyar & Morozov, 1976); Staccato articulation (Juslin, 1997); strong amplitude and the harmonic content stronger than in sadness and weaker than in anger (Salgado, 2000); fast and light vibrato, and sharp contrast between long and short notes (Gabrielsson & Juslin, 1996).

Facially: In this condition and according to the research done on the subject (see validation of the stimuli in Experiment IV), the involved muscles are: *Zygomaticus Major* and *Zygomaticus Minor*, *Risorius*, *Orbicularis Oculi* and *Frontalis* (see Ekman's example in the figure below):

Ekman's example photograph of a facial expression of emotion 'happiness'

Interpretation - Fear

This emotional peak corresponded to Toni's attempt to express vocally and facially the emotion of fear (bar 23 and 24). This moment occurred around the 35th second of the performance.

Vocally: In this condition, the vocal cues involved should be: Variable high and low sound level (Juslin, 2000); fast mean tempo (Juslin, 1997); fast tone attacks (Kotlyar & Morozov, 1976); Staccato articulation (Juslin, 1997); Timbre – many harmonics (Scherer & Osinsky, 1977); strong amplitude and the harmonic content stronger than in sadness and weaker than in anger (Salgado, 2000); very fast vibrato and variation in sound intensity (Gabrielsson & Juslin, 1996).

Facially: In this condition the muscles *Levator Labii Superioris, Frontalis (pars medialis), Orbicularis Oculi, Caninus,* and *Levator Palpebrae Superioris* should be involved (see Ekman's example in the figure below):

Ekman's example photograph of a facial expression of emotion 'fear'

Interpretation – Sadness

This emotional peak corresponded to Toni's attempt to express vocally and facially the emotion of sadness (bar 34 and 35) and this moment occurs around the 50th second of the performance.

Vocally: In this condition, the vocal cues involved should be: Low sound level (Juslin, 2000); dull timbre (Juslin, 2000); slow mean tempo (Gabrielsson & Lindström, 1995); slow tone attacks (Kotlyar & Morozov, 1976); Legato articulation (Juslin, 1997); Timbre - few harmonics (Scherer & Oshinsky, 1977); low amplitude and weak harmonic content (Salgado, 2000).

Facially: In this condition the muscles *Corrugator, Triangularis, Depressor Supercilii, Mentalis* and *Depressor Labii Inferioris,* should be used extensively (see Ekman's example in the figure below):

Ekman's example photograph of a facial expression of Emotion 'sadness'

Interpretation - Anger

The outburst of the traveller's anger corresponded to Toni's attempt to express vocally and facially this emotion (bars 44, 45 and 46), and this moment occurred around the 70th second of the performance.

Vocally: In this condition, the vocal cues involved should be: High sound level (Juslin 2000), sharp timbre (Juslin, 2000); fast mean tempo (Scherer & Oshinsky, 1977); abrupt tone attacks (Kotlyar & Morozov, 1976); accents on unstable notes (Lindström, 1999); Timbre - many harmonics (Scherer & Oshinsky, 1977); dramatic change in both amplitude and harmonic content (Salgado, 2000).

Facially: In this condition the involved muscles should be: *Platysma, Procerus, Depressor Labii Inferioris, Mentalis, Levator Labii Superioris,* and *Aleque Nasi* (see Ekman's example in the figure below):

Ekman's example photograph of a facial expression of Emotion 'anger'

Summary

To conclude, it has been stated that along the development of the musical structure of the song 'Die Post', four different musical moments correspond to four distinctive emotional categories. Happiness, fear, sadness and anger are musically translated and introduced by Schubert in a dialogue between emotion and reason, which contribute to carry out a work of unequalled poetry and emotional strength. It also has been examined how the two levels of the expressive musical code, the 'introversive structural/dynamic coding' and the 'extroversive-structural/dynamic coding' (see Chapter 3, pg. 111 and 123), work and inter-relate, and how the interpreter could corroborate the expressive means of the musical structure through the use of facial and vocal expressive performance cues (Indexical-qualitative coding).

structure through the use of facial and vocal expressive performance cues (Indexical-qualitative coding).

Due to the appropriate musical structures, melodies, harmonies, rhythms, timbres and textures this song brings the richness and the most intimate of human experience and of the human soul. The interpreter should be able to rediscover the emotions presented through the musical structures and to communicate them to the audience, through the corroborating expressive cues. Of course, in all the vocal and facial effects attention should be paid to not interfere with the necessary singing technical adjustments. Nevertheless, and according to the results obtained in Experiment III, it seems that for the exception of the action of the muscle Platysma, involved in the expression of anger, no secondary adjustments must be made in the expression of the four emotional meanings considered in the interpretation of this song.

This structural and interpretative analysis of the Lied *Die Post* was used for the preparation of the performance of the Schubert's song within the context of Experiment IV, and later on within the PhD final recital. It assisted the performer to communicate the song's musical meaning to the audience in a clear understandable way. Of course, the accuracy of the performer's intended emotional expressiveness will be finally achieved through the use of a practical tool, which will assist the improvement and definition of the facial and vocal gestures (the expressive performance cues) required to express the analysed emotional meaning of the musical structure. For this reason the study undertaken on Experiment-IV, aiding the elaboration of the expressive performance tool, might be of relevance for the final recital and for the performance practice of singing. It might aid to assist performers to improve the communication and artistry of their performances, and it might be an important guide and tool for singing teachers and students of singing for the learning of the mechanisms of expressiveness accuracy and meaning communication in music performance.

D. SCORE OF THE SCHUBERT SONG: *DIE POST.*

mein___ Herz?_____ Die

Post bringt kei-nen Brief für dich. Was drängst___ du denn so wun-der-

fp *fp* *pp*

lich, mein Herz,___ mein Herz?_____ Die

pp *p*

Post bringt kei-nen Brief für dich, mein Herz,___ mein Herz,___ was drängst du denn so

cresc. *p*

wun-derlich, mein Herz,_____ mein Herz?___

f *p*

Nun ja, die Post kommt aus der Stadt, wo

ich ein lie - bes Lieb-chen hatt, mein Herz!

cresc. *f*

wo ich ein lie - bes Lieb-chen hatt,

p *decresc.* *pp*

mein Herz, _____ mein_ Herz!_____

fp *fp*

1

Willst wohl ein-mal hin-ü-ber-sehn und fra - gen,

wie es dort mag gehn, mein Herz, mein Herz?

willst wohl ein-mal hin - ü - ber-sehn, mein Herz, mein

Herz, und fra - gen, wie es dort mag gehn, mein Herz,

mein Herz?

Appendix G: Software's Programme Text (Experiment IV)

EMOTIONS IN VIDEOS

Beginning of Application
Define the video's list to be used
Start with a video type "Verbal answer" – "Sound On – Video Off"

```
on exitFrame me
  global list_video, in_video, type
 list_video =[]
  repeat with a=1 to 10
    add (list_video, member (a, "video").name)
  end repeat
  in_video=1
  type=1

  put the short date&& the long time&return&"Verbal
  answer"&return&return&"Sound On - Video Off"&return into field
  "user_data"
end
```

Video start
Starts with an aleatory video, but not repeated, type active
Verifies if it is a type of video "Sound On – Video Off" – if it is prepare the histogram of the sound to show

```
global list_video, in_video, type, turn, ttotal, gopen, actualv,
retry, type

on beginSprite me

  if retry=1 then
    sprite(the currentspritenum).member=actualv

    if type =1 or type =4 then
      sprite(25).loch=400
      sprite(25).member=actualv&"s"
      ttotal=member(actualv).duration
      the timer=0
      sprite(26).loch=sprite(25).left
      gopen=1
    else
      gopen=0
    end if

    case type of
      1:
        member (the member of sprite(the currentspritenum)).sound=1
        member (the member of sprite(the currentspritenum)).video=0
      2:
        member (the member of sprite(the currentspritenum)).sound=0
```

```
          member (the member of sprite(the currentspritenum)).video=1
      3:
        member (the member of sprite(the currentspritenum)).video=1
        member (the member of sprite(the currentspritenum)).sound=1
      4:
        member (the member of sprite(the currentspritenum)).video=0
      5:
        member (the member of sprite(the currentspritenum)).sound=0
        member (the member of sprite(the currentspritenum)).video=1
      6:
        member (the member of sprite(the currentspritenum)).video=1
        member (the member of sprite(the currentspritenum)).sound=1
    end case
    retry-2
  else
    retry=0
    n=getat (list_video, random(count(list_video)))
    actualv=n
    types-type
    sprite(the currentspritenum).member=n
    if typo =1 or typo =4 then
      sprite(25).loch=400
      sprite(25).member=n&"s"
      ttotal=member(n).duration
      the timer=0
      sprite(26).loch=sprite(25).left
      gopen=1
    else
      gopen=0
    end if

    put n&return after field "user_data"
    deleteone (list_video, n)

    ---choose type of video part to show
    case type of
      1:
        member (the member of sprite(the currentspritenum)).sound=1
        member (the member of sprite(the currentspritenum)).video=0
      2:
        member (the member of sprite(the currentspritenum)).sound=0
        member (the member of sprite(the currentspritenum)).video=1
      3:
        member (the member of sprite(the currentspritenum)).video=1
        member (the member of sprite(the currentspritenum)).sound=1
      4:
        member (the member of sprite(the currentspritenum)).video=0
      5:
        member (the member of sprite(the currentspritenum)).sound-0
        member (the member of sprite(the currentspritenum)).video=1
      6:
        member (the member of sprite(the currentspritenum)).video=1
        member (the member of sprite(the currentspritenum)).sound=1
    end case


    if in video=10 then
      in_video=1
      list_video=[]
```

```
      repeat with a=1 to 10
        add (list_video, member (a, "video").name)
      end repeat
      type=type+1
      turn=1
    else
      in_video=in_video+1
    end if
  end if
end
```

Emotion Choice

Emotion Choice

Register of emotion choice and of answer's timing

```
global retry

on mouseup
  retry=0
  the floatPrecision = 2
  put (the timer/60.0)&&"seg."&&member(the member of sprite(the
currentspritenum)).name&return after field "user_data"
  the timer=0
  go "scale"
end
```

Emotion differential

Register of emotion differential and of answer's timing

Verify if change of Video type is necessary , for instance change to "Sound On – Video On"

```
global type, turn
on mouseUp me
  the floatPrecision = 2
  put (the timer/60.0)&&"seg."&&"Scale:"&& integer(the
currentspritenum-6)&return&return after field "user_data"

  if type <>8 then
    if turn=1 then
      case tipo of
        1: put "Sound On - Video Off"&return after field "user_data"
        2: put "Sound Off - Video On"&return after field "user_data"
        3: put "Sound On - Video On"&return after field "user_data"
        4: put "Non-verbal answer"&return&return&"Sound On - Video
Off"&return after field "user_data"
        5: put "Sound Off - Video On"&return after field "user_data"
        6: put "Sound On - Video On"&return after field "user_data"
        7:
          put "Sound Track"&return after field "user_data"
          go "excerpt"
          abort
      end case
      turn=0
    end if
    go "v1"
  end if
end
```

Final answer

Register of emotion and of answer's timing

Sees the final answer, prepares a name to the file with data and saves the file with the user's answers

```
on  mouseup
  the floatPrecision = 2
  put (the timer/60.0)&&"seg."&&member(the member of sprite(the
currentspritenum)).name&return after field "user_data"

  if the frame<45 then
    go "excerpt2"
    abort
  end if


  --Prepare a name  to the file
  a=the short date&"_"&the short time
  a=(the short date.char [1..2])&"_"&(the short date.char
[4..5])&"_"&(the short date.char [7..10])
  b=""
  if the short time.char [2]<>":" then
    a=a&"_"&(the short time.char [1..2])&"_"&(the short time.char
[4..5])
  else
    a=a&"_"&(the short time.char [1])&"_"&(the short time.char [3..4])
  end if
  nameregister=a

  --------- SAVE IN A FILE RTF-------------------
  put field "user_data" into print

  set myFile = new(xtra "fileio")
  createFile( myfile, "c:\logs\"&a&".rtf" )
  openFile(myFile, "c:\logs\"&a&".rtf", 0)
  writeString( myfile, print )
  set theFile = readFile(myFile)
  closeFile(myFile)
  set myFile = 0

  go "end"
end
```

PERFORMANCE

MANAGER

Beginning of the aplication

Start variables
Creates a file to control start and end of the performance

```
on exitFrame me
  global step, review

  step=0
  review=0

  steps="Pause"
  ---create a field
  myFile = new(xtra "fileio")
  filenames=the moviepath&"register.sal"
  createFile( myfile, filenames )
  openFile(myFile, fileNames, 0)
  writeString( myfile, steps)
  theFile = readFile(myFile)
  closeFile(myFile)
  myFile = 0
end
```

Beginning of the performance

Indicates that performance is beginning

```
global step

on mouseUp me
  step=integer(step+1)
  steps-"step"&step

  member ("steps").text=string(step)
  ---create a field

  repeat with a-1 to 8
    myFile - new(xtra "fileio")
    filenames=the moviepath&"register.sal"
    -- createFile( myfile, filenames )
    openFile(myFile, fileNames, 0)
    writeString( myfile, steps)
    theFile = readFile(myFile)
    closeFile(myFile)
    myFile = 0
  end repeat

  repeat with a=200 to 800
    puppetsprite a, 0
  end repeat
  go "generate"
end

on beginSprite me
  if passo=2 then sprite(the currentspritenum).loch=-200
end
```

End of the performance

Performance ends and indicates to the users to stop register

Goes into the generation of the graph

```
Global step
on mouseUp me
  if step=2 then steps="finit"
  else ateps="Pause"

  repeat with a=1 to 8
    myFile = new(xtra "fileio")
    filenames=the moviepath&"registo.sal"
    -- createFile( myfile, filenames )
    openFile(myFile, fileNames, 0)
    writeString( myfile, stepss)
    theFile = readFile(myFile)
    closeFile(myFile)
    myFile - 0
  end repeat
  the timer=0
  repeat while the timer<100
  end repeat

  go to "graph final"
end
```

Generate the graph with the answers of the 3 posts

Takes the information from the text fields of the users

Generates the graph based on that information

```
global step, nline, review, nline1, nline2, nline3

on exitFrame me
  --reed the fields and generate graph
  repeat with a=1 to 3
    myFile = new(xtra "fileio")
    filenames=the moviepath&"user"&a&".txt"
    openFile(myFile, filenames, 1)
    if status(myFile) = 0 then
      theFile = readFile(myFile)
      put theFile into field "datauser"&a
    else
      alert error(myfile,status(myfile))
    end if
    closeFile(myFile)
    myFile = 0
  end repeat

  happiness=0
  fear=0
  sadness=0
  anger=0

  --already taken
```

```
total_line1=member("datauser1").line.count
total_line2=member("datauser2").line.count
total_line3=member("datauser3").line.count

if review<>0 then
  case review of
    1:first_line1=4
      first_line2=4
      first_line3=4

      total_line1=nlines1
      total_line2=nlines2
      total_line3=nlines3

    2:first_line1=nlines1
      first_line2=nlines2
      first_line3=nlines3

    3:first_line1-4
      first_line2=4
      first_line3=4
  end case
else
  if step=2 then
    first_line1=nlines1
    first_line2=nlines2
    first_line3=nlines3
  else
    first_line1=4
    first_line2=4
    first_line3=4
    nlines1=total_line1
    nlines2=total_line2
    nlines3=total_line3
  end if
end if

total1=(member("datauser1").line[total_line1-
4].word[3])+(member("datauser1").line[total_line1-2].word[3])
  total2=(member("datauser2").line[total_line2-
4].word[3])+(member("datauser2").line[total_line2-2].word[3])
  total3=(member("datauser3").line[total_line3-
4].word[3])+(member("datauser3").line[total_line3-2].word[3])
  the floatPrecision = 2
  total=max(total1, total2, total3)

  the floatPrecision = 2
  divide=(total/670.00)

  repeat with t=1 to 10
    member("t"&t).text=string((total/10*t)/60)
  end repeat
  repeat with nuser=1 to 3
    name_user="datauser"&nuser
    case nuser of
      1: hgh=first_line1
        hgh1=total_line1
      2: hgh=first_line2
        hgh1=total_line2
```

```
      3: hgh=first_line3
         hgh1=total_line3
    end case
    repeat with a=hgh to hgh1
      if member(name_user).line[a].word[1] = "Emotion:" then
    --    alert string (a)
        sensa=member(name_user).line[a].word[2]
        timeelapsed=member(name_user).line[a-1].word[3]
        gpressure=member(name_user).line[a+1].word[3]
        case sensa Of
          "happiness", "happyness":
            posv=130
            happiness=happiness+1
          "fear":posv-190
            fear=fear+1
          "sadness":posv=70
            sadness=sadness+1
          "anger":posv=250
            anger-anger+1
          "neutral":posv=0
        end case
        nsprite=(a+(20*nuser*10))
        put nsprite
        puppetsprite nsprite, true
        sprite (nsprite).ink=36
        sprite (nsprite).member="traço_base"&nuser
        sprite(nsprite).width=integer(gpressure/divides)+1
        sprite(nsprite).height=10
        sprite(nsprite).loch=integer(timeelapsed/divides)+95
        sprite(nsprite).backcolor=nuser
        sprite(nsprite).locv=(posv+nuser*15)
        -- sprite(nsprite).width=integer(gpressure/divides)

        member("cliques"&nuser).line[1]=happiness
        member("cliques"&nuser).line[2]=fear
        member("cliques"&nuser).line[3]=anger
        member("cliques"&nuser).line[4]=sadness
        updatestage
      end if
    end repeat
  end repeat
end
```

USER

Choice of post

Choose post and place to save the registers

Create a file to save the registers

```
global paths, npost, pathgeral

on exitFrame me
  ---open window asking path
  set myFile = new(xtra "fileio")
  if the machinetype = 256 then
    setfiltermask (myfile, "Choice of file,*.sal")
  else
    setfiltermask (myfile, " Choice of file, *.sal")
  end if
  fileNames = displayOpen(myFile)
  pathgeral=filenames
  if not voidP(filenames) and not (filenames = EMPTY) then
    if status(myFile) - 0 then
      member("path").text=filenames
      bb=filenames.length-11
      paths=filenames.char[1..bb]
    else
      alert error(myfile,status(myfile)) -show message of error
    end if
  end if
  closeFile(myFile)                          -- Close file
  set myFile = 0                             -- Turn off the instance

  ---create field
  myFile = new(xtra "fileio")
  filenames=paths&"user"&nposto&".txt"
  createFile( myfile, filenames )
  openFile(myFile, fileNames, 0)
  writeString( myfile, member("dados user").text )
  theFile = readFile(myFile)
  closeFile(myFile)
  myFile = 0
end
```

Choice of buttons

Register of the buttons to use for each emotion

```
global tanger, thappiness, tsadness, tfear
on keyup
  member("whichbutton").text=the key

  case the frame of
    7: thappiness=the keycode
    8:tfear=the keycode
    10:tanger=the keycode
    11:tsadness=the keycode
  end case
  updatestage
end
```

Control of the session

Registes of the emotions

when session ends, keep all registers in a text field

```
global pathgeneral, gstep, gbaset, paths, npost
global tanger, thappiness, tneutral, tsadness, tfear
on exitFrame me
  go the frame
  verify_alter
end

on keydown
  gbaset=the timer
end

on keyup
  if gstep="pause" then abort
  if gstep<>"step1" and gstep<>"step2" then abort
  case the keycode of
    tanger: emot="Anger"
    thappiness: emot="Happyness"
    tsadness: emot="Sadness"
    tmedo: emot="Fear"
    otherwise
      abort
  end case
  the floatPrecision = 2
  put "Elapsed time:"&&(the
timer)&return&"Emotion:"&&emoc&return&"Pressure time:"&&((the timer-
gbaset))&return&return after field "data user"
end

on verify_alter
  myFile = new(xtra "fileio")
  openFile(myFile, pathgeneral, 1)
  -- if status(myFile) = 0 then
  theFile = readFile(myFile)
  put theFile into field "instruct"
  -- else
  --    alert error(myfile,status(myfile))
  --    end if
  closeFile(myFile)
  myFile = 0

    --- reacts according to the state of the session
  if gpasso=member("instruct").text then abort

  if member("instruct").text="pause" then
    gstep="pause"
    gpause=1
    sprite(3).member="inpause"

    myFile = new(xtra "fileio")
    filenames=paths&"user"&nposto&".txt"
    -- createFile( myfile, filenames )
    openFile(myFile, fileNames, 0)
    writeString( myfile, member("data user").text )
```

```
  theFile = readFile(myFile)
  closeFile(myFile)
  myFile = 0

  abort
end if
if member("instruct").text="finit" then
  myFile = new(xtra "fileio")
  filenames=paths&"user"&nposto&".txt"
  -- createFile( myfile, filenames )
  openFile(myFile, fileNames, 0)
  writeString( myfile, member("datauser").text )
  theFile = readFile(myFile)
  closeFile(myFile)
  myFile = 0

  go movie "enter"
  abort
end if
gpause=0
sprite(3).member="session"
gstep=field "instruct"
the timer=0

--writes
---save in a field
myFile = new(xtra "fileio")
filenames=paths&"user"&npost&".txt"
-- createFile( myfile, filenames )
openFile(myFile, fileNames, 0)
writeString( myfile, member("datauser").text )
theFile = readFile(myFile)
closeFile(myFile)
myFile = 0

  put gstep&return after field "datauser"
end
```

REFERENCES AND BIBLIOGRAPHY

REFERENCES AND BIBLIOGRAPHY

Adachi, M. & Trehub, S. (1998). Children's expression of emotion in song. *Psychology of Music,* **26**, 133-53.

Amstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the nature of language.* Cambridge: Cambridge University Press.

Bach, C.P.E. (1778). *Essay on the true art of playing keyboard instruments,* (Ed. and tr. W. J. Mitchell). London: Eulenburg Books, 1985. (Originally published in 1778).

Baddeley, A. & Logie, R. (1992). Auditory imagery and working memory. In D. Reisberg (Ed.), *Auditory imagery,* pp. 179-97. Hillsdale, NJ: Lawrence Erlbaum.

Baily, J. (1977). Movement Patterns in Playing the Herati Dutàr. In J. Blacking (Ed.), *The Anthropology of the Body,* pp. 215-330. London: Academic.

Baily, J. (1985). Music structure and human movement. In P. Howell, I. Cross, and R. West (Eds.), *Musical Structure and Cognition,* pp. 337-58. London: Academic Press.

Balk, H.W. (1985). *Performing power: A new approach for the singer-actor.* Minneapolis: University of Minnesota Press.

Balkwill, L-L. & Thompson, W.F. (1999). A cross-cultural investigation of the perception of Emotion in music: psychophysical and cultural cues. *Music Perception,* **17**, 43-64.

Balzano, G.J. (1987). Measuring Music. In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music,* pp. 177-99. Stockholm: Royal Swedish Academy of Music.

Banse, R. and Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology,* **70**, N° 3, 614-636.

Barker, A. (1989). *Greek Musical Writings, Vol. II, Harmonic and Acoustical Theory.* Cambridge, MA: Cambridge University Press.

Baroni, M. & Finarelli, L. (1994). Emotions in spoken language and vocal music. In I. Deliège (Ed.), *Proceedings of the Third International Conference for Music Perception and Cognition,* pp. 343-5, Liége, Belgium: University of Liége.

Barthes, R. (1985). Rasch. *The Responsibility of Forms: Critical Essays on Music, Art, and Representation,* (tr. R. Howard), pp. 299-312. NY: Hill and Wang.

Barthes, R. (1989). One Always Fails in Speaking of What One Loves. In *The Rustle of Language* (tr. R. Howard), pp. 296-305. Berkeley: University of California Press.

Bartlet, D.L. (1996). Physiological responses to music and sound stimuli. In D. A. Hodges (Ed.), *Handbook of music psychology,* (2nd edn), pp. 348-85. San Antonio, TX: IMR.

Bassilli, J.N. (1978). Facial motion in the perception of faces and emotional expression. *Journal of Personality and Social Psychology,* **4**, 373-379.

Bassilli, J.N. (1979). Emotion recognition: The role of facial movement. *Journal of Personality and Social Psychology*, 37, 2049-58.

Bauer, M. (1915). *Die Lieder Franz Schuberts*. Leipzig.

Beardsley, M.C. (1966). *Aesthetics From Classical Greece to the Present: A Short History*. Tuscaloosa: University of Alabama Press Beardsley.

Beaufils, M. (1982). *Le Lied Romantique Allemand*. Éditions Gallimard.

Becking, G. (1975). *Über ein dänisches Schul-Liederbuch, über Mitbewegung und Gehatsanalyse*. In E. Kramolisch (Ed.), pp. 191-218. Originally in *Zeitschrift für Musikwissenschaft* (1923-4).

Bezooijen, R. (1984). *The characteristics and recognizability of vocal expression of emotions*. Dorderecht, The Netherlands: Foris.

Bierwisch, M. (1979). Musik und Sprache: Überlegungen zu ihrer Structur und Funktionsweise. In E. Klemm (Ed.), *Jahrbuch Peters 1978*, pp. 9-102 Leipzig: Edition Peters.

Bloothooft, G. (1985). *Spectrum and timbre of sung vowels*. PhD diss., Vrije Universiteit te Amsterdam.

Boer, G. H. (1992). How might emotions affect learning? In S.-A. Christianson (Ed.), *The handbook of emotion and memory: research and theory*, pp. 3-31. Hillsdale, NJ: Erlbaum.

Borod, J.C., Obler, L., Erhan, H.M., Gruwald, I.S., Cicero, B.A., Welkowitz, J. et al. (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology*, 12, 446-58.

Bowman, W. D. (1998). *Philosophical Perspectives on Music*, Oxford: Oxford University Press.

Bregman, A.S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge: Cambridge University Press.

Bruce, V. & Young, A. (1998). *In the Eye of the Beholder: The Science of Face Perception*. Oxford, NY: Oxford University Press.

Brunswick, E. (1956). *Perception and representative design of psychological experiments*. Berkeley, CA: University of California Press.

Bryden, M.P., Ley, R. & Sugerman, J. (1982). A left-ear advantage for identifying the emotional quality of tonal sequences. *Neuropsychologia*, 20, 83-7.

Budd, M. (1985). *Music and the Emotions: The Philosophical Theories*. London: Routledge & Kegan Paul.

Bull, P. (1983). *Body movement and interpersonal communication*. Chichester: John Wiley & Sons.

Bunt, L. & Pavlicevic, M. (2001). Music and emotion: perspectives from music therapy. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, 10, 181-201. Oxford University Press.

Burrows, D.L. (1990). *Sound, Speech, and Music*. Amherst: University of Massachusetts Press.

Caccini, G. (1614). *Le nuove musiche e nuove manere di scriverle*. Reprinted facsimile, Florence: Sudio per eddizione Scelte, 1983.

Cacioppo, J.T., Klein, D.J., Bernston, G.G., & Hatfield, E. (1993). The psychophysiology of Emotion. In M. Lewis & J. Havilland (Eds.), *Handbook of emotions*, pp. 119-142. New York: Guilford.

Calder, A.J., Young, A.W. Perret, D.I., Etcoff, N. L. and Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition*, 3, 81-117.

Callen, D. (1983). Transfiguring Emotions in Music. *Philosophische Grazier Studien*, 16, 69-91.

Capell, R. (1957). *Schubert's Songs*. London.

Chomsky, N. (1957). *Syntactic structures*. Mouton: The Hague.

Clarke, E. (2001). Meaning and the specification of motion in music. In *Musicae Scientiae*, Fall, vol. V, 2, 213-234.

Clarke, E. & Davidson, J. W. (1998). The body in performance. In W. Thomas (Ed.) *Composition, Performance, Reception: Studies in the Creative Process in Music*, pp. 74-92. Ashgate Publishing Company.

Clarke, E. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross and R. West (Eds.), *Musical Structure and Cognition*, pp. 209-36. London: Academic Press.

Clarke, E. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in Music*, pp. 1-26. Oxford: Clarendon,.

Clarke, E. (1995). Expression in performance: generativity, perception and semiosis. In J. Rink (Ed.), *The Practice of Performance*, pp. 21-54. Cambridge University Press.

Clifton, T. (1983). *Music as heard: A Study in Applied Phenomenology*. New Haven CT: Yale University Press.

Clynes, E. F. (1969). Precision of essentic form in living communication. In K.M. Leibovic and J. C. Clynes (Eds.), *Information Processing in the Nervous System*. New York: Springer-Verlag.

Clynes, E. F. (1973). Sentics: Biocybernetics of emotion communication. *Annals of the New York Academy of Sciences*, 220, 55-131.

Clynes, E. F. (1977). *Sentics: The touch of emotions*. New York: Doubleday.

Clynes, M. (1980). Transforming emotionally expressive touch to similarly expressive sound. In *Proceedings Tenth International Acoustic Congress, Sydney.*

Clynes, M. (1986). Generative principles of musical thought. Integration of microstructure with structure. *Communication and Cognition AI*, 3, 185-223.

Cohen, M.R. (1944). *A Preface to Logic*. New York: H. Holt & Co.

Cook, N. & Dibben, N. (2001). Musicological approaches to emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 45-70.Oxford: University Press.

Cook, N. (1987). *A Guide to Musical Analysis*. London: J.M.Dent & Sons Ltd.

Cook, N. (1995). The conductor and the theorist: Furtwängler, Schenker and the first movement of Beethoven's Ninth Symphony. In J. Rink (Ed.), *The Practice of Performance: Studies in Musical Interpretation*. Cambridge: Cambridge University Press.

Cooke, D. (1959). *The language of music*. London: Oxford University Press.

Cooksey, R.W. (1996). *Judgement analysis*. New York: Academic Press.

Cox, A. (2001). The mimetic hypothesis and embodied musical meaning. *Musicae Scientiae*, Fall, vol. V, 2, 195-212.

Crowder, R.G. (1985). Perception of major/minor distinction: III. Hedonic, musical, and affective descrimination. *Bulletin of the Psychonomic Society*, 23, 314-16.

Cutting, J.E. (1982). Two ecological perspectives: Gibson vs. Shaw and Turvey. *American Journal of Psychology*, 95, 199-222.

Cutting, J.E., & Kozlowski, L.T. (1977). Recognising friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353-56.

Damasio, A. (1994). *Descartes'error: Emotion, reason, and the human brain*. New York: Avon Books.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Davidson, J.W. & Correia, J.S. (2001). Meaningful Musical Performance: A bodily experience. *Research Studies in Music Education*, 17, 70-83.

Davidson, J.W. & Correia, J.S. (2002). Body Movement. In R. Parncutt and G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*, pp. 237-50. Oxford: Oxford University Press.

Davidson, J.W. (1991). *The Perception of Expressive Movement in Music Performance*. Unpublished Doctoral Dissertation. London: City University, London.

Davidson, J.W. (1993). Visual perception of performance manner in movements of solo musicians. In *Psychology of Music*, 21/2, 103-13.

Davidson, J.W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies*, 26, 279-301.

Davidson, J.W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. In *Musicae Scientiae*, Fall, vol. V, 2, 235-256.

Davies, S. (1994). Musical meaning and expression. Ithaca, NY: Cornell University Press.

Desain, P. & Honing, H. (1992). *Music, Mind and Machine*. Amsterdam: Thesis Publishers.

Desain, P. & Honing, H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research*, 56, 285-292.

Dewey, J. (1984). The Theory of Emotion. *Psychological Review*, I, 553-69.

Dolging, K. & Adelson, E. (1990). Age changes in the ability to interpret in sung and instrumentally-presented melodies. *Psychology of Music*, 18, 87-98.

Dommel-Diény, A. (1976). *L'Analyse Harmonique en Exemples. Schubert-Liszt*. Ed. A. Dommel-Diény.

Dowling, W. J. & Harwood, D.L. (1986). *Music Cognition*. San Diego: Academic Press.

Draffan, K.R. (2000). *Singing of the Soul: An investigation of the impact of self-concept, confidence, Emotion, and social environment on the Singer's voice and ability to perform*. Unpublished MMus Dissertation, University of Sheffield.

Duchenne, G.-B. (1872). *Mechanisme de la physionomie humaine ou analyse électro-physiologique de l'expression des passions applicable a la pratique des arts plastiques*. Paris: Renouard.

Eco, U. (1972). *La Structure absente*. Paris: Mercure de France.

Eisenberg, N. & Strayer, J. (1987). *Empathy and its development*. Cambridge, UK: Cambridge University Press.

Ekman, P. & Davidson, R.J. (Eds.) (1994). *The nature of Emotion: Fundamental questions*. New York: Oxford University Press.

Ekman, P. & Friesen, W.V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P. & Friesen, W.V. (1978). *Manual for the facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P. (1984). Expression and the nature of the emotion. In K. R. Scherer and P. Ekman (Eds.), *Approaches to emotion*, pp. 319-343. Hillsdale, NJ: Erlbaum.

Ekman, P. (1989). The argument and evidence about universals in facial expressions of emotion. In H. Wagner and A. Manstead (Eds.), Handbook of social psychophysiology, 58, 342-353.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.

Ekman, P. (1993). Facial expression and Emotion. *The American Psychologist*, 48, 384-392.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russel's mistaken critique. *Psychological Bulletin*, 115, 268-287.

Ekman, P., Friesen, W.V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings.* New York: Pergamon Press.

Ekman, P., Levenson, R.W., & Friesen, W.V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science,* **221,** 1208-10.

Eliot, D.J. (1987). Structure and Feeling in Jazz: Rethinking Philosophical Foundations. In *Bulletin of the Council for Research in Music Education,* **95,** 7-26.

Elliot, R.K. (1966-7) Aesthetic Theory and the Experience of Art. *Proceedings of the Aristotelian Society,* 67, 111-126.

Epstein, D. (1981). On musical continuity. In J. T. Fraser and N. Lawrence (Eds.), *The Study of Time,* vol. IV, 180-97. New York: Springer-Verlag.

Epstein, D. (1993). On Affect and musical motion. In S. Feder, R. Karmel, and G. Pollock (Eds.), *Psychoanalitic Explorations in Music,* pp. 91-123. Madison, Connecticut: International Universities Press, Inc.

Fadiga, L. and Gallese, V. (1997). Action representation and language in the brain. In *Theoretical Linguistics,* **23,** 267-80.

Fadiga, L., Buccino, G. Craighero, L. Fogassi, L., Gallese, V. & Pavesi, G. (1998). Corticospinal excitability is specifically modulated by motor imagery: a magnetic stimulation study. In *Neuropsychologia,* **37,** 147-58.

Fairbanks, G. (1940). Recent experimental investigation of vocal pitch in voice. *Journal of the Acoustical Society of America,* **11,** 457-66.

Faulkner, R.S.C. (1999), *Expression in Performance.* Submitted essay in partial fulfilment MA in Psychology for Musicians, 7. University of Sheffield.

Feil, A. (1975). *Franz Schubert: Die Schöne Müllerin, Winterreise.* Stuttgart: Philip Reclam

Feld, S. (1984). Communication, Music, and Speech about Music. *Yearbook for Traditional Music,* **16,** 1-18.

Feldman, J., Epstein, D., & Richards, W. (1992). Force dynamics of tempo change in music. *Music Perception,* **10,** 185-204.

Fischer-Dieskau, D. (1971). *Auf den Spuren der Schubert-Lieder.* Wiesbaden

Födermayr, F. and Deutsch, W.A. (1994). Parimi veder le lagrime. One aria, three interpretations. In A. Friberg, J. Iwarsson, E. Jansson, and J. Sundberg (Eds.), *Proceedings of the Stockholm Music Acoustics Conference, July 1993,* 96-107. Stocholm: Publications issued by the Royal Sweedish Academy of Music.

Fonagy, I. (1962). Mimik auf glottaler Ebener. In *Phonetica,* **8,** 209-219.

Fonagy, I. (1967). Hörbare Mimik. In *Phonetica,* **16,** 25-35.

Fonagy, I. (1976). La mimique buccale. In *Phonetica,* **33,** 31-44.

Fonagy, I., & Magdics, K. (1976). Emotional patterns in intonation and music. In *Zeitung fuer Phonetik*, 16, 293-326.

Fowler, C.A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3-28.

Frances, R. (1958). *La perception de la musique*, Paris, Vrin.

Friberg, A. & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final *ritardandi* derived from measurements of stopping runners. Journal of the *Acoustical society of America*, 105, 1469-1484.

Frick, R. (1985). Communicating Emotion: the role of prosodic features. *Psychological Bulletin*, 97, 412-429

Frijda, N.H. (1986). *The emotions*. Cambridge: Cambridge Press University.

Frijda, N.H. (1994). Emotions are functional, most of the time. In P. Ekman, & R. J. Davidson (Eds.), *The nature of Emotion: Fundamental questions*, 112-122. New York: Oxford University Press.

Frijda, N.H. (2000). The psychologist's point of view. In M. Lewis, & J.M. Havilland-Jones (Eds.), *Handbook of emotions* (4ᵗʰ ed.), pp. 59-74. New York: The Guilford Press.

Gabrielsson, A. & Juslin, P. (1996). Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. In *Psychology of Music*, 24, 68-91.

Gabrielsson, A. & Lindström, E. (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology*, 14, 94-116.

Gabrielsson, A. & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, 10, 223-248. Oxford University Press.

Gabrielsson, A. (1973a). Similarity ratings and dimension analyses of auditory patterns, I and II. *Scandinavian Journal of Psychology*, 14, 138-76.

Gabrielsson, A. (1973b). Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology*, 14, 244-60.

Gabrielsson, A. (1982). Perception and Performance of Musical Rhythm. In M. Clynes (Ed.), *Music, Mind, and Brain*. New York: Plenum Press.

Gabrielsson, A. (1988) Timing in music performance and its relations to music experience. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition*, pp. 27-51. Oxford: Clarendon Press

Gabrielsson, A. (1991). Experiencing music. *Canadian Journal of Research in Music Education*, 33, 21-6.

Gabrielsson, A. (1994). Intention and emotional expression in music performance. In A. Friberg, J. Iwardson, E. Jansson, & J. Sundberg (Eds.), *Proceedings of the Stockholm Music Acoustics Conference.*

Gabrielsson, A. (1995). Expressive Intention and Performance. In R. Steinberg (Ed.), *Music and the Mind Machine: The psychophysiology and psychopathology of the sense of music,* pp. 35-47. Berlin, Heidelberg, New York, London: Springer-Verlag.

Gabrielsson, A. (1999). The performance of music. In D. Deutchs (Ed.), *The Psychology of Music* (2nd ed.), pp. 501-602, San Diego, CA: Academic Press.

Gabrielsson, A. (2001). Emotions in strong experiences with music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research,* 19, 431-449. Oxford: Oxford University Press.

Galile, V. (1581), *Dialogo della musica antica e della moderna.* Reprinted facsimile: In O. Strunk (Ed.), Source Reading in Music History (1950). New York: W. W. Norton.

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mid-reading. In *Trends in Cognitive Sciences,* 2, 493-501.

Gerardi, G.M. & Gerken, L. (1995). The Development of affective responses to madality and melodic contour. *Music Perception,* 12, 279-90.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1973). On the concept of "formless" invariants in visual perception. *Leonardo,* 6, 43-45. Reprinted in E. S. Reed and R. K. Jones (Eds.), *Reasons for Realism: selected essays of James J. Gibson* (1982). Erlbaum Associates.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Gibson, J.J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology,* 49, 182-194.

Gibson, J.J. (1960). The information contained in light. *Acta Psychologica,* 17, 23-30.

Gibson, J.J., & Gibson, E.J. (1955). Perceptual Learning: Differentiation or enrichment? *Psychological Review,* 62, 32-41.

Goblot, E. (1901). La musique descriptive. *Revue philosophique,* LII, 58-77.

Goldstein, A. (1980). Thrills in response to music and other stimuli. *Physiological Psychology,* 8, 126-9.

Greymas, A.J. & Courtés, J. (1979). *Sémiotique: Dictionnaire raisonné de la Théorie du Language.* Paris: Hachette.

Gundlach, R.H. (1935). Factors determining the characterization of musical phrases. *American Journal of Psychology,* 47, 624-44.

Hahn, R. (1920). *Du Chant.* Paris.

Handel, S. (1991). *Listening. An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Hanslick, E. (1957). *The Beautiful in Music.* Reprinted in Morris Weitz (Ed.) (trans. Gustav Cohen) (7th ed. of 1885). New York: Liberal Arts Press. (1st ed. 1854).

Hanslick, E. (1986). *On the musically beautiful: A contribution towards the revision of the aesthetics of music* (trans. G. Payzant). Indianopolis: In Hackett. (*Vom Musikalisch-Schönen,* originally published 1854).

Hatfield, E. Cacioppo, J.T. & Rapson, R.L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Hatten, R.S. (1999). *Lectures on Musical Gesture, 1-8.* Cyber Semiotic Institute Home Page: Course outline.

Hawkes, T. (1977). *Structuralism and Semiotics.* London: Methuen.

Helmholtz, H.L.F. von [1863] (1954). *On the sensations of a tone as a psychological basis for the theory of music.* New York: Dover

Hemsley, T. (1998). *Singing and Imagination.* Oxford University Press.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology,* **48**, 248-68.

Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology,* **49**, 621-30.

Hillman, J. (1960). *Emotion: A Comprehensive Phenomenology of Theories and Their Meanings for Therapy.* London: Routledge and Kegan Paul.

Hjelmslev, L. (1971). *Prolégomènes à une théorie du langage.* (1st ed. 1943). Paris: Minuit.

Hjortsjö, C. H. (1969*) Man's face and mimic language.* Studentlitteratur, Malmö, Sweden.

Hodges, D. & Haack, P. (1999) In D. A. Hodges (ed.), *Handbook of music psychology.* IMR Press: University of Texas at San Antonio.

Holt, J. (1996) A comprehensivist theory of art. *The British Journal of Aesthetics,* **36**, 424-431.

Howes, F. (1958) *Music and Its Meanings.* London: Athlone Press.

Husserl, E. (1964). *Phenomenology of Internal Time-Consciousness.* In M. Heidegger (Ed.), (tr. J. C. Churchill). Bloomington: Indiana University Press.

Husson, R. (1962) *Le Chant.* Paris: Presses Universitaires de France.

Imberty, M. (1975). Perspectives nouvelles pour l'étude expérimentale des associations verbales à l'oeuvre musicale. *Musique en Jeu,* **17**, 87-109.

Imberty, M. (1979). *Entendre la musique.* Paris: Dunod.

Imberty, M. (1997). Can one seriously speak of narrativity in music? In A. Gabrielsson (Ed.), *Proceedings of the third Triennial ESCOM Conference, 1977,* pp. 13-22. Uppsala, Sweden: Department of Psychology, Uppsala University.

Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Izard, C.E. (1977). *The emotions*. New York: Plenum Press.

Izard, C.E. (1980). Cross-cultural perspectives on emotion and emotion communication. In H. Triandis & W. J. Lonner (Ed.), *Handbook of cross-cultural psychology*, Vol. 3, pp. 63-75. Boston: Allyn & Bacon

Izard, C.E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, 99, 561-565.

Izard, C.E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. In *Psychological Bulletin*, 115, 288-299.

Jackendoff, R. (1988). *Consciousness and the Computational Mind*. Cambridge, Mass: MIT Press.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press, 1987.

Johnstone, T., & Scherer, K.R. (2000). Vocal communication of Emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed.), pp. 220-235. New York: The Guildford Press.

Jowett, B. (1990). Excerpt where otherwise indicated all references to Plato are drawn from *The Dialogues of Plato* (tr. B. Jowet), in *Great Books of the Western World*. Chicago: Encyclopaedia Britannica.

Jürgens, U. & von Cramon, D. (1982). On the role of the anterior cingulate cortex in phonation: A case report. *Brain and Language*, 15, 234-48.

Juslin P.N. & Laukka P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae*, Fall 2000, Vol. IV, 2, 151-183.

Juslin P.N. & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception*, 17, 197-221.

Juslin, P.N. & Persson, R.S. (2002). Emotional Communication. In R. Parncutt and G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*, pp. 219-236. Oxford: Oxford University Press.

Juslin, P.N. (1995). Emotional communication in music viewed through a Brunswikian lens. In *Musical expression*. Proceedings of the conference of ESCOM and DGM 1995, in G. Kleinen (Ed.), pp. 21-5. Bremen, Germany: University of Bremen.

Juslin, P.N. (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception*, 14, 383-418.

Juslin, P.N. (1998). Functionalist perspective on emotional communication in music performance. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 78*, pp. 35-51. Uppsala, Sweeden: Uppsala University Library.

Juslin, P.N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1797-813.

Juslin, P.N. (2001). Communicating emotion in music performance: a review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 309-337. Oxford: Oxford University Press.

Kainz, F. (1962). *Psychologie der Sprache. I. Band. Grundlagen der allgemeinen Sprchspsychologie*, 3rd ed. Stuttgart: Enke.

Kappas, A., Hess, U., & Scherer, K.R. (1991). Voice and Emotion. In R.S. Feldman and B. Rime (Eds.), *Fundamentals of Nonverbal Behavior: Studies in Emotion and Social Interaction*. Cambridge: Cambridge University Press.

Kendall, R.A. & Carterette, E.C. (1990). The Communication of musical expression. *Music Perception*, 8, 129-164.

Kerman, J. (1985). *Contemplating Music: Challenges to Musicology*. Cambridge, Mass.: Harvard University Press.

Kivy, P. (1980). *The Corded Shell: Reflections on Music Expression*. Princeton, NJ: Princeton University Press.

Kivy, P. (1989). *Sound sentiment*. Philadelphia, PA: Temple University Press.

Kleinen, G. (1968). *Experimentelle Studien zum musikalischen Ausdruck*. Hamburg, Germany: Universität Hamburg.

Knepler, G. (1977). *Geschichte als Weg zum Musikverständnis: Zur Theorie, Methode und Geschichte der Musikgeschichtsschreibung*. Leipzig: Reclam.

Kotlyar, G. M. & Morozov, V. P. (1976). Acoustical correlates of the emotional content of vocalised speech. *Soviet Physiology and Acoustics*, 22, 208-211.

Kozlowski, L.T. & Cutting, J. E. (1977). Recognising the sex of a walker from a dynamic point-light display. In *Perception and Psychophysics*, 21, 575-80.

Kramer, L. (1984). *Music and Poetry. The Nineteenth Century and After*. Berkeley, LA, London: University of California Press.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music*, 21, 3-19.

Kristeva, J. (1989). *Language the Unknown: An Initiation into Linguistics* (tr. A. M. Menke). New York: Columbia University Press.

Kronman, U., & Sundberg, J. (1987). Is the musical ritard an allusion to physical motion? In A. Gabrielsson (Ed.), *Action and perception in rhythm and music*. Stockholm: Publications issued by the Royal Swedish Academy of Music, 55, 57-68.

Krumhansl, C.L. (1996). A perceptual analysis of Mozart's Piano sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, 13, 401-32.

Krumhansl, C.L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, 51, 336-52.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University Press.

Langeheinecke, E.J., Schnitzler, H.-U-, Hischer-Buhrmeister, M., & Behene, K.-E. (1999). Emotions in singing voice: Acoustic cues for joy, fear, anger, and sadness. Poster presented at the Joint Meeting of the Acoustical Society of America and the Acoustical Society of Germany, Berlin, March 1999.

Langer, S. K. (1953). *Feeling and Form: A theory of Art Developed from Philosophy in a New Key*, New York: Charles Scribner's Sons.

Langer, S.K. (1942). *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*, Cambridge: Harvard University Press. [(1951). 2nd edn., New York: New American Lybrary; and (1957). 3rd edn., Cambridge, MA: Harward University Press.].

Lashley, K.S. (1951). The problem of serial order in behaviour. In L. A. Jeffries (Ed.), *Cerebral mechanisms in behaviour*, pp. 1-18. New York: Willey.

Lavy, M.M. (2001). *Emotion and the Experience of Listening to Music: A Framework for Empirical Research*. Unpublished Doctoral Dissertation. Cambridge: Jesus College,

Lazarus, R.S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lechte, J. (1994). *Fifty Key Contemporary Thinkers: From Structuralism to Postmodernity*. NY, London: Routledge.

LeDoux, J. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Touchstone.

Lerdal, F. & Jackendorf, R. (1983). *A generative theory of tonal music*. Cambridge, Massachusstes: MIT Press.

Lewis, M. & Haviland-Jones, J.M. (Eds.). (1993). *Handbook of emotions*. New York: The Guildford Press.

Lidov, D. (1987). Mind and body in music. *Semiotica*, 66, 69-97.

Lieberman, P. & Michaels, S.B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *JASA*, 34, 922-27.

Lindström, E. (1997). Impact of melodic structure on emotional expression. In A. Gabrielsson (Ed.), *Proceedings of the third triennial ESCOM Conference, Uppsala, June 1997*, pp. 292-7. Uppsala, Sweeden: Uppsala University.

Lindström, E. (1999). Expression in music: Interaction between performance and melodic structure. Paper presented at the Meeting of the Society for Music Perception and Cognition, Evanston, USA, 14-17 August 1999.

Lipps, Th. (1909). *Leitfaden der Psychologie*. Leipzig, Germany: Engelmann.

Lochead, J. (1995). Hearing New Music: Pedagogy from a Phenomenological Perspective. *Philosophy of Music Education Review*, 3,1-36.

Lomax, A. (1968) *Folk song style and culture*. New Brunswick, NJ: Transaction Books.

Lorenz, K. (1965). *Evolution and modification of behaviour*. Chicago: University Press.

Lundqvist, L.G., Carlsson, F., & Hilmersson, P. (2000). Facial electromyography, autonomic activity, and emotional experience to happy and sad music. Paper presented at the 27th International Congress of Psychology, Stockholm, Sweden, 23-28 July 2000.

Madsen, C.K. & Fredrickson, W.E. (1993). The experience of musical tension: A replication of Nielsen's research using the continuous response digital interface. *Journal of Music Therapy*, 30, 46-63.

Madsen, C.K. (1997). Emotional responses to music. *Psychomusicology*, 16, 59-67.

Maher, T.F. (1980). A rigorous test of the proposition that musical intervals have different psychological effects. *American Journal of Psychology*, 93, 308-27.

Maher, T.F., & Berline, D.E. (1982). Verbal and exploratory responses to melodic music intervals. Psychology of Music, 10, 11-27.

Manén, L. (1974). *The Art of Singing*. Faber Music Ltd.

Margolis, J. (1987). On the Semiotics of Music. In Ph. Alperson (Ed.), *What is Music?*, pp. 211-36. NY: Haven Press.

Mattheson, J. (1739). *Der Vollkommene Capellmeister* (tr. E. C. Harris). Ann Arbor, MI: UMI Research Press, 1981.

McIntosh, D.N. (1996). Facial feedback hypotheses: Evidence, implications, and directions. *Motivation and Emotion*, 20, 121-47.

Mead, G.H. (1934). *Mind, self, and Society*. Chicago: University of Chicago Press.

Merleau-Ponty, M. (1962). *Phenomenology of Perception*, (tr. C. Smith). London; Routledge and Kegan Paul.

Merleau-Ponty, M. (1964). *The Primacy of Perception and Other Essays on Phenomenological Psychology, the Philosophy of Art History, and Politics*. In J.M. Edie (Ed.). Evanston, IL: Northwestern University Press.

Merriam, A. (1964) *The Anthropology of Music*. Chicago: Northwestern University Press

Meyer, L.B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Meyer, L.B. (1967). *Music, the Arts, and Ideas: Patterns and Predictions in Twentieth-Century Culture*. Chicago: University of Chicago Press.

Meyer, L.B. (1973). *Explaining Music*. Los Angeles: University of California Press.

Meyer, L.B. (1994). Emotion and Meaning in Music. In R. Aiello (Ed.), with J. A. Sloboda, *Musical perceptions*, pp. 3-37. New York: Oxford University Press.

Meyer, L.B. (2001). Music and emotion: distinctions and uncertainties. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 341-60. Oxford University Press.

Middleton, R. (1990). *Studying Popular Music.* Milton Keynes: Open University Press.

Molino, J. (1978). Fait musical et sémiologie de la musique, *Musique en jeu*, **17**, 37-62.

Nakamura, T. (1987). The communication of dynamics between musicians and listeners through musical performance. *Perception & Psychophysics*, **41**, 525-533.

Nattiez, J.J., (1975). *Fondements d'une sémiologie de la musique.* Paris: Union générale d'éditions.

Nattiez, J.J., (1990). *Music and Discourse: Toward a Semiology of music*, (tr. C. Abbade).

Neumann, R. & Strack, F. (2000). Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, **79**, 211-23.

Newham, P. (1998) *Therapeutic Voice Work.* J. Kingsley Publishers.

Nielsen, F.V. (1983). *Experience of musical tension.* Copenhagen: Akademisk Forlag.

Nielsen, F.V. (1987). Musical 'tension' and related concepts. In T. A. Sebeok & J. Umiker-Sebeok (Ed.), *The semiotic web '86. An international yearbook*, pp. 491-513. Berlin: Mouton de Gruyter.

Nielzén, s. & Cesarec, Z. (1981). On the perception of emotional meaning in music. *Psychology of Music*, **9**, 17-31.

Nielzén, s. & Cesarec, Z. (1982). Emotional experience of music as a function of musical structure, *Psychology of Music*, **10**, 7-17.

Noy, P. (1993). How music conveys Emotion. In S. Feder, R. Karmel, & G. Pollock (eds.), *Psychoanalitic Explorations in Music, 125-149.* Madison, Connecticut: International Universities Press, Inc.

Nyklicek, I., Thayer, J.K., & van Doornen, L.J.P. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, **11**, 304-21.

Oatley, K. (1992). *Best laid schemes. The psychology of emotions.* Cambridge, MA: Harvard University Press.

Ohgushi, K. & Hattori, M. (1996a). Emotional Communication in Performance of Vocal Music. In B. Pennycook & E. Costa-Giomi (Ed.), *Proceedings of the 4th ICMPC*, pp. 269-274. Montreal, Canada: McGill University.

Ohgushi, K. & Hattori, M. (1996b). Acoustic correlates of the emotional expression in vocal performance. Paper presented at the *Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*, Honolulu, Hawaii, 2-6 December 1996.

Orlov, H. (1981). Toward a Semiotics of Music. In W. Steiner. Austin (Ed.), *The Sign in Music and Literature*, pp. 131-137. University of Texas Press.

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *The Behavioral and Brain Sciences*, 5, 407-467.

Panksepp, J. (1995). The emotional sources of 'chills' induced by music. *Music Perception*, 13, 171-208.

Papealiou, C. & Trevarthen, C. (1998). The infancy of music. *Musical praxis*, 1, 19-33.

Papousek, H. (1996). Musicality in infancy research: biological and cultural origins of early musicality. In I. Deliège & J. Sloboda (Eds.), *Musical Beginnings: Origins and development of musical competence*, pp. 88-112. Oxford/New York/Tokyo: Oxford University Press.

Papousek, M. (1996). Intuitive parenting: a hidden source of musical stimulation in infancy. In I. Deliège & J. A. Sloboda (Eds.), *Musical Beginnings: Origins and development of musical competence*, pp. 88-112. Oxford: Oxford University Press.

Patel, A.D. & Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In I. Deliége & J. A. Sloboda (Ed.), *Perception and cognition of music*, pp. 191-215. Hove, UK: Psychology Press.

Pavlov, I. (1927). *Conditional Reflexes*. London: Oxford University Press.

Peck, S.R. (1951) *Atlas of Human Anatomy for the Artist*. New York: Oxford University Press.

Peck, S.R. (1987) *Atlas of Facial Expression*. New York: Oxford University Press.

Peirce, C.S. (1931-38). *Collected Papers of Charles Sanders Peirce*. In C. Hartshorne, P. Weiss, & A. W. Burk (Eds.), 8 vols. Cambridge, Mass: Harvard University Press.

Peretz, I. (2001). Listen to the brain: a biological perspective on musical emotions. In P. Juslin & J. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 105-134. Oxford: Oxford University Press.

Peretz, I., Gagnon, L. & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68, 111-41.

Persson, R.S. (1995). Musical reality: Exploring the subjective world of performers. In R. Monelle & C. T. Gray (Ed.), *Song and signification. Studies in music semiotics*, pp. 58-63. Edinburgh, UK: Faculty of Music, University of Edinburgh.

Persson, R.S. (2001). The subjective world of the performer. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 275-289. Oxford: Oxford University Press.

Persson, R.S., Pratt, G., & Robson, C. (1992). Motivational and influential components of musical performance: A qualitative analysis. *European Journal for High Ability*, 3, 206-217.

Pierce, A. (1978). Structure and phrase, part I. *In Theory Only*, 4, 22-35.

Pierce, A. (1994). Developping Schenkerian Hearing and Performing. *Intégral*, 8, 51-123.

Plato. *The Dialogues of Plato* (translated by B. Jowett, 1952). In *Great Books of the Western World*, Vol. 1. Chicago: Encyclopaedia Britannica.

Ploog, D. (1986). The biological foundations of the vocal expressions of emotions. In . R. Plutchik & H. Kellerman *Emotion* (Ed.), *Theory, research, and experience. Vol. 3: Biological foundations of emotion*, pp. 173-97. New York: Academic Press.

Plutchik, R. (1994). *The psychology and biology of emotion.* New York: Harper-Collins.

Plutchik, R., & Kellerman, H. (Eds.). (1989). *Emotion: Theory, research, and experience. Volume 4: The Measurement of Emotions.* San Diego: Academic Press Inc.

Pratt, C.C. (1931). *The Meaning of Music: A Study in Psychological Aesthetics.* New York: McGraw-Hill.

Pratt, C.C. (1954). The Design of Music. *Journal of Aesthetics and Art Criticism*, 12, 289-300.

Princeton, NJ: Princeton, University Press.

Proust, M. (1934). *Remembrance of things past.* Vol. 1, (tr. C. K. Scott Moncrieff). New York: Random House.

Rapoport, D. (1950). *Emotions and Memory.* New York: International Universities Press Inc.

Rapoport, E. (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research*, 25, 109-49.

Rasch, R.A. (1988). Timing and synchronisation in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition*, pp. 27-51. Oxford: Clarendon Press

Reber, A.S. (1985). *The Penguin dictionary of psychology.* New York: Penguin Books.

Reed, E.S. & Jones, R.K. (1982). *Reasons for Realism: Selected Essays of James J. Gibson.* Erlbaum.

Reed, E.S. (1988). *James J. Gibson and the Psychology of Perception.* Yale University Press.

Repp, B. (1992). Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's Träumerei. *Journal of the Acoustical Society of America*, 92, 2546-568.

Repp, B. (1993). Music as motion: a synopsis of Alexader Truslit's (1938) "Gestaltung und Bewegung in der Musik". *Psychology of Music*, 21, 48-72.

Repp, B. (1994). On determining the basic tempo of an expressive music performance. *Psychology of Music*, 22, 157-167.

Rigg, M.G. (1937). Musical expression: An investigation of the theories of Erich Sorantin. *Journal of Experimental Psychology*, 21, 442-55.

Rigg, M.G. (1939). *What features of a musical phrase have emotional suggestiveness?* Publications of the Social Science Research Council of the Oklahoma Agricultural and Mechanical College, 1.

Rigg, M.G. (1940). The effect of register and tonality upon musical mood. *American Journal of Psychology*, 2, 49-61.

Rigg, M.G. (1964). The mood effects of music: A comparison of data from four investigators. *American Journal of Psychology*, 58, 427-38.

Rink, J. (1990). Musical Structure and Performance. *Music Analysis*, 9, 319-39.

Robson, C. (1993). *Real World research. A resource for social scientists and practitioner-researchers.* Oxford, UK: Blackwell.

Rosch, R. (1978). Principles of categorisation, In E. Rosch & B.B.Loyd (eds.).*Cognition and Categorisation*, pp. 27-48. Hillsdale, NJ: Erlbaum.

Rosolato, G. (1974). La voix: entre le corps et langage. *Revue francaise de psychanalyse*, vol. 38, 1, 75-94.

Rousseau, J.J. [1761] (1986). Essay on the origin of Languages. In J. H. Moran & A. Gode (Eds.), *On the origin of language: Two essays*, pp. 5-74. Chicago: University of Chicago Press.

Runeson, S. and Frykholm, G. (1983). Kinematic Specification of Dynamics as an informational basis for person-and-action perception: Expectation, gender, recognition, and deceptive intention. In *Journal of Experimental Psychology: General*, 112, 585-615.

Russel, J.A. & Fernandez-Dols, J.M. (Eds.). (1997). *The psychology of facial expression.* Cambridge, UK: Cambridge University Press.

Russel, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-78.

Ruwet, N. (1967). Musicologie et linguistique. *Revue Internationale des Sciences Sociales*, vol. XIX, 1, 85-93.

Ruwet, N. (1972). *Language, musique e poésie.* Paris: Le Seuil.

Sachs, C. (1943). *The Rise of Music in the Ancient World East and West.* London: Dent & Sons.

Salgado, A. (2000). Voice, emotion and facial expression in singing. In *Proceedings of 6th International Conference on Music Perception and Cognition, Keel, UK, 5 - 9 de August 2000.*

Salgado, A. (2001). Contribution to the understanding of some of the processes involved in the perception and recognition of emotional meaning on singers' facial and vocal expression. In *Proceedings of the 1st Meeting of the Argentina Society for the Cognitive Sciences of Music, Buenos Aires, Mai 2001.*

Salgado, A. (2002a). Measuring facial e-motion in singing. In *Proceedings of the Royal College of Music Conference, London, April 2002.*

Salgado, A. (2002b). Investigating emotional meaning in music performance: singing. In *Proceedings of the 7*[Th] *International Conference on Music Perception and Cognition, Sydney, Australia, August 2002.*

Sallis, J. (1973). *Phenomenology and the Return to Beginnings.* Pittsburgh: Duquesne University Press.

Salmon, P.G. & Meyer, R.G. (1992). *Notes from the green room: Coping with stress and anxiety in musical performance.* San Francisco: Jossey-Bass Publishers.

Saussure, F. (1966). *Course in General Linguistics,* (tr. W. Baskin). New York: McGraw Hill.

Schenker, H. (1935). *Der Freie Satz,* (tr. E. Oster, 1979). New York: Longman.

Scherer, K., Wherle, T., Kaiser, S., & Schmidt, S. (2000). Studying the Dynamics of Emotional Expression Using Synthesised Facial Muscle Movements. *Journal of Personality and Social Psychology,* Vol. 78, 1, 105-119.

Scherer, K.R. & Oshinsky, J.S. (1977). Cue utilisation in emotion attribution from auditory stimuli. *Motivation and Emotion,* I, 331-46.

Scherer, K.R. & Siegwart, H. (1995). Acoustic Concomitants of Emotional Expression in Operatic Singing: The Case of Lucia in *Ardi gli incensi. Journal of the Voice,* vol. 9, 3, 249-260. Philadelphia: Lippincott-Raven Pub.

Scherer, K.R. & Zentner, M.R. (2001). Emotional effects of music: production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research,* pp. 361-392. Oxford University Press.

Scherer, K.R. (1980). The functions of nonverbal signs in conversation. In R. St. Clair, H.Gilles (Eds.), *The social and psychological contexts of language,* pp. 225-44. Hillsdale, NJ: Erlbaum,

Scherer, K.R. (1981). Speech and emotional states. In J. K. Darby (Ed.), *Speech evaluation in psychiatry,* pp. 189-220. New York: Grune & Stratton.

Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin,* 99, 143-65.

Scherer, K.R. (1991). Emotion expression in speech and music. In J. Sundberg, L. Nord, & L. Carlson (Eds.), *Music, language, speech and brain,* pp. 146-156. London: Macmillan

Scherer, K.R. (1992). What does facial expression express? In Strongman (Ed.), *International review of studies on Emotion,* 2, 139-65. Chichester, UK: Willey.

Scherer, K.R. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion,* 7, 1-41.

Scherer, K.R. (1994). Affect bursts. In van S.H.M. Goozen, van de N.E. Poll, & J.A. Sergeant (Eds.), *Emotions: essays in Emotion theory,* pp. 161-196. Hillsdale, NJ: Erlbaum.

Scherer, K.R. (1995). Expression of Emotion in Voice and Singing. *Journal of Voice*, vol. 9, 3, 235-248.

Schlosberg, H. (1941). A scale for the judgement of facial expressions. *Journal of Experimental Psychology*, 29, 497-510.

Schmalfeldt, J. (1985). On the Relation of Analysis to Performance: Beethoven's Bagatelles Op. 126, N° 2 and 5. *Journal of Music Theory*, 29, 1-31.

Schmidt, L. & Trainor, L. (2001). *Frontal brain electrical activity (EGG) distinguishes valence and intensity of musical emotions.* Manuscript submitted for publication (quoted in Peretz, 2001:119).

Schubert, E. (1999). *Measurement and time series analysis of emotion in music.* Unpublished doctoral dissertation, University of south Wales, Sydney, Australia.

Schubert, E. (2001). Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 393-414. Oxford: Oxford University Press.

Scruton, R. (1997). *The Aesthetics of Music.* Oxford: Oxford University Press.

Seashore, H.G. (1937). An objective analysis of artistic singing. In C. E. Seashore (Ed.), *Objective analysis of musical performance: University of Iowa studies in the psychology of music*, vol. 4, pp. 12-157. Iowa City, IA: University of Iowa.

Sedlacek, K., & Sychra, A. (1963). Die Melodie als Faktor des emotionellen Ausdrucks. *FP*, 15, 89-98.

Senju, M. Ohgushi, K. (1987). How are the player's ideas conveyed to the audience? *Music Perception*, 4, 311-24.

Shaffer, L.H. (1995). Musical Performance as interpretation. *Psychology of Music*, 23, 17-38.

Shaffer, L.H. (1984). Timing in musical performance. In J. Gibbon & L. Allan (Eds.), *Timing and time perception*. Annals of the New York Academy of Sciences, 423, 420-28.

Shaffer, L.H. (1989). Cognition and affect in musical performance. *Contemporary Review of Music*, 4, 381-389.

Shaffer, L.H. (1992). How to interpret Music. In M. R. Jones and S. Holleran (Eds.), *Cognitive Basis of Musical Communication*, pp. 263-78. Washington: American Psychological Association.

Shaver, P.R., Schwartz, J., Kirson, d., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52, 1061-86.

Shaw, R. & Pittenger, J. (1978). On perceiving change. In H. Pick and E. Saltzman (eds.), *Modes of Perceiving and Processing Information*. Hillsdale, New Jersey: Erlbaum, 187-204.

Shepherd, J. (1991). *Music as a Social Text.* Cambridge: Polity Press.

Shepherd, J. and Wicke P. (1997). *Music and Cultural Theory*, Cambridge: Polity Press.

Shove, P. & Repp, B.H. (1995), Musical motion and performance: theoretical and empirical perspectives. In J. Rink (Ed.), *The Practice of Performance. Studies in Musical Interpretation*, pp. 55-83. Cambridge: Cambridge University Press.

Sievers, E. (1924). *Ziele und Wege der Schallanalyse*, Heidelberg: Carl Winter's Universitätsbuchhandlung.

Sloboda, J.A. & Juslin, P.N. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research*, pp. 71-104. Oxford: Oxford University Press.

Sloboda, J.A. (1983) The communication of musical metre in piano performance. *Quaterly Journal of Experimental Psychology*, **35A**, 377-396.

Sloboda, J.A. (1985) *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Oxford University Press.

Sloboda, J.A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, **19**, 110-20.

Sloboda, J.A. (1992). Empirical studies of emotional response to music. In M.Riess-Jones & S. Holleran (Eds.), *Cognitive bases of musical communication*, pp. 33-46. Washington, DC: American Psychological Association.

Sloboda, J.A. (1996a). The acquisition of musical performance expertise: Deconstructing the 'talent' account of individual differences in musical expressivity. In K. A. Ericsson (Ed.), *In the road to excellence*, pp. 107-26. Mahwah, NJ: Erlbaum.

Sloboda, J.A. (1996b). Emotional responses to music: A review. In K. Riederer & T. Lahti (Ed.), *Proceedings of the Nordic Acoustical Meeting*, pp. 385-92. Helsinki: The Acoustical Society of Finland.

Sloboda, J.A. (1998). Does music mean anything? *Musicae Scientiae*, Spring, Vol. II, 1, 21-31.

Sloboda, J.A. (2000). Musical performance and Emotion: Issues and developments. In S. W. Yi (ed.), Music, Mind, and Science, pp. 220-238. Seoul, Korea: Western Music Research Institute.

Sloboda, J.A. (Ed.) (1988) *Generative processes in music: The psychology of performance, improvisation, and composition*. Oxford: Clarendon Press.

Sloboda, J.A., Lehmann, A.C., & Parncutt, R. (1997). Perceiving intended Emotion in concert-standard performances of Chopin's Prelude N° 4 in E-minor. In A. Gabrielsson (Ed.), *Proceedings of the Third Triennial ESCOM Conference*, pp. 629-34. Uppsala, Sweden: Department of Psychology, Uppsala University.

Small, C. (1987). Performance as ritual: sketch for an inquiry into the true nature of a symphony concert. In A. L. White (Ed.), *Lost in Music: Culture, Style and the Musical Event*. Hanover, London: Routledge & Kegan Paul.

Small, C. (1998). *Musicking: The Meanings of Performing and Listening.* Hanover, London: Wesleyan University Press.

Smith, J.A. (1995). Semi-structured interviews. In J. A. Smith, R. Harré, & L. V. Langenhove (Eds.), *Rethinking Methods in Psychology,* pp. 9-26. London: Sage Publications.

Snook, S. (1998). *Kinesthetic analysis in post-tonal instrumental composition.* Unpublished doctoral dissertation, University of Sydney.

Spitz, R.A. (1963). Ontogenis: The proleptic function of emotion. In P. H. Knapp (Ed.), *Expressions of the emotions in man,* pp. 36-60. New York: International Universities Press.

Stern, D.N. (1985). *The interpersonal world of the infant: A view from the psychoanalysis and developmental psychology.* New York: Basic Books.

Strauss, C. (1971). Bolero de Maurice Ravel. *L'Homme,* Vol. 11, 2, 5-14.

Strauss, C. (1972). *Structural Anthropology,* (trans. C. Jacobson & B. G. Schoepf). Harmondsworth: Penguin.

Strongman, K.T. (1996). *The psychology of Emotion: Theories of Emotion in perspective.* Chichester: John Wiley & Sons.

Stubley, E. (1966). Modulating Identities: The Play of Ensemble Performance (Being in the Body, Being in the Sound). Paper presented to the Music Educators National Conference. Kansas City, MO.

Sundberg, J. & Verrillo, V. (1980). On the anatomy of the ritard: a study of timing in music. *Journal of the Acoustical Society of America,* 68, 772-79.

Sundberg, J. (1974). Articulatory interpretation of the "singing" formant. *JASA,* 55, 838-44.

Sundberg, J. (1982). Speech, song and emotions. In M. Clynes (Ed.), *Music, Mind, and Brain,* pp. 137-149. New York: Plenum Press.

Sundberg, J. (1987). *The Science of the Singing Voice.* Illinois: Northern Illinois University Press. (Translation of "Röstlära" by Johan Sundberg, Proprius Förlag, Stockholm, 1980).

Sundberg, J. (1988). Computer synthesis of music performance. In J. A. Sloboda (Ed.), *Generative processes in Music,* 52-69. Oxford: Clarendon.

Sundberg, J. (1999). The Perception of Singing. In D. Deutsch (Ed.), *The Psychology of Music,* pp. 171-214. San Diego, California: Academic Press.

Sundberg, J., Frydén, L., & Friberg, A. (1995). Expressive aspects of instrumental and vocal performance. In R. Steinberg (Ed.), *Music and the Mind Machine: The psychophysiology and psychopathology of the sense of music,* pp. 49-62. Berlin, Heidelberg, New York, London: Springer-Verlag.

Sundberg, J., Iwarsson, J., & Hagegard, H. (1995). A singer's expression of emotions in sung performance. In O. Fujimura & M. Hirano (Ed.), *Vocal fold physiology: Voice quality control,* pp. 217-29. San Diego, CA: Singular Press.

Swanwick, K. (1988). Musical Development: the first years. In *Music, Mind and Education,* IV, pp. 36-54. London: Routledge.

Tagg, P. (1991). *Fernando the Flute: Analysis of Musical Meaning in an ABBA Mega-Hit.* Liverpool: The Institute of Popular Music, University of Liverpool.

Thompson, W.F. & Robitaille, B. (1992). Can composers express emotions trough music? *Empirical Studies of the Arts,* 10, 79-89.

Tiffin, J. & Seashore, H.G. (1932). Summary of the established facts in experimental studies on the vibrato up to 1932. In C. E. Seashore (Ed.), *University of Iowa studies in the psychology of music: vol. I. The vibrato,* pp. 344-376. Iowa City: University of Iowa.

Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception,* 3, 33-58.

Todd, N. P. (1989). A computational model of rubato. *Contemporary Music Review,* 3, 69-88.

Todd, N. P. (1992). The dynamics of dynamics: a model of musical expression. *Journal of the Acoustical Society of America,* 91, 3450-550.

Todd, N. P. (1995). The kinematics of musical expression. In *Journal of the Acoustical Society of America,* 97, 1940-949.

Todd, N. P. (1999). Motion in music: a neurobiological perspective. *Music Perception,* 17, 115-26.

Tomkins, S.S. (1962). *Affect, imagery, and consciousness: The positive affects.* New York: Springer.

Tomkins, S.S. (1984). Affect theory. In K. R. Scherer & P. Ekman (Ed.), *Approaches to Emotion,* pp. 163-96. Hillsdale, NJ: Erlbaum.

Trehub, S.E., Unyk, A.M., & Trainor, L.J. (1993). Adults identify infant-directed music across cultures. *Infant Behaviour and Development,* 16, 193-211.

Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientia,* Special Issue, 115-211.

Truslit, A. (1938). *Gestaltung und Bewegung in der Musik.* Berlin-Lichterfelde: Chr. Friedrich Vieweg.

Vaneechoute, M. & Skoyles, J. (1998). The memetic origin of language: Modern humans as musical primates. *Journal of Memetics: Evolutionary Models of Information Transmission,* 2, 84-117.

Wagner, C. (1974). Experimentelle Untersuchungen über das Tempo. *Österreichische Musikzeitschrift,* 29, 589-604.

Wagner, H. (1989). The peripheral physiological differentiation of emotions. In H. Wagner and A. Manstead (Eds.), *Handbook of psychophysiology: Emotion and social behaviour,* pp. 63-75. Chichester: Wiley and Sons.

Walton, K. (1997). Listening with imagination: is music representational? In J. Robinson (Ed.), *Music and meaning,* pp. 57-82. Ithaca: Cornell University Press.

Watson, K.B. (1942). The nature and measurement of musical meanings. *Psychological Monographs*, **54**, 1-43.

Wedin, L. (1972). Multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, **13**, 241-57.

Williams, C.E. & Stevens, K.N. (1972). Emotions and speech: Some acoustic correlates. *JASA*, **52**, 1238-50.

Witvliet, C.V., Vrana, S.R., & Webb-Talmage, N. (1998). In the mood: Emotion and facial expressions during and after instrumental music, and during an emotional inhibition task. *Psychophysiology Supplement*, **88**.

Woodworth, R.S. (1938). *Experimental psychology*. New York: Holt.

Wundt, W. (1900). Völkerpsychologie. *Eine untersuchung der Entwicklungsgesetze von sprache, Mythos und Sitte*. Leipzig: Kröner.

Wundt, W. (1987). *Outlines of psychology* (trans. C. H. Judd). Leipzig: Englemann.

Zebrowitz, L.A. (1990). *Social Perception*. Open University Press.

Zentner, M.R., Meylan, S., & Scherer, K.R. (2000). Exploring 'musical emotions' across five genres of music. Paper presented at the 6th International Conference of the Society for music Perception and Cognition (ICMPC), 5-10 August, Keele, UK.