The Sonic Dimension of the Psychoanalytic Encounter: An Interdisciplinary Study Integrating Psychoanalysis, Music Theory and Developmental Research

Thesis submitted for the degree of “Doctor of Philosophy”

By

Adva Frank-Schwebel

Submitted to the Senate of the Hebrew University of Jerusalem

October, 2010
Contents

Chapter I – Introduction

I.1 Subject and Goal of the Research.............................................................. 3
I.2 Description and Characteristics of the Nonlinguistic Sonic Dimension........... 10
I.3 Sound and Music in Psychoanalysis.......................................................... 16
I.4 The Justification for an Interdisciplinary Research..................................... 23
I.5 A Conceptual Context for the Investigation of
the Nonlinguistic Sonic Dimension............................................................... 29
I.6 Chapter Outline.................................................................................... 34

Chapter II - A Conceptual Context: Sound in Phenomenology, Music Semiotics, and
Dialogic Theory

II.1 Introduction.......................................................................................... 37
II.2 A Phenomenological Approach to Sound.................................................. 38
II.3 Sonic Meaning and Signification: Sound in music semiotics...................... 60
II.4 Dialogism and Sound............................................................................ 83
II.5 Theoretical Interrelations...................................................................... 96
II.6 Developmental Research and its Relation to the Proposed Conceptual
Context....................................................................................................... 105

Chapter III – The Voice in Psychoanalysis

III.1 Introduction.......................................................................................... 108
III.2 A Critical Discussion of the Voice in Psychoanalysis: Three Psychoanalytic
Hearings....................................................................................................... 112
III.3 A Critique of Psychoanalytic Hearings of the Voice.................................. 129
III.4 A Multidimensional Hearing of the Voice in Psychoanalysis.................... 142
Chapter IV - Listening

IV.1 Introduction.......................................................................................................................... 233
IV.2 Psychoanalytic Listening, Past and Present................................................................. 235
IV.3 The Musical Nature of the Auditory Experience in the Clinical Situation....... 240
IV. 4 Listening - in a Manner of Conclusion.................................................................................260

References........................................................................................................................................269
Chapter I
Introduction

1.1 Subject and Goal of the Research

This study is an investigation of the nonlinguistic sonic dimension and its role in the clinical setting of psychoanalysis and psychoanalytic psychotherapy. Its main premise is the idea that nonlinguistic vocal expressions are implicitly perceived and organized by the analytic dyad\(^1\) into a system of communication, creating a sonic dimension that lies mostly outside of immediate awareness. The nonlinguistic sonic dimension has an essential role in creating a semiotic field that exceeds propositional meanings and unfolds through different mechanisms than the verbal, having several functions. Primarily, it establishes the therapeutic relation by simultaneously creating and reflecting the unfolding of the interpersonal, intersubjective matrix. This is achieved through the representation and negotiation of interpersonal forces by vocal means. Secondly, the sonic dimension articulates the subjectivity of each participant through the signification and representation of their affective and energetic\(^2\) states in the course of conversation. These functions are embodied, performative and presentational, i.e. inherent in the lived, present moment of the analytic dialogue. Furthermore, it is proposed that the sonic dimension can be conceived of in psychoanalytic terms, in that it can represent certain unconscious ideation.

\(^1\) The term “analytic dyad” refers to the patient and the therapist in the psychoanalytic or psychotherapeutic setting, without making any theoretic assumption about the nature of the relationship. In addition, the terms “analytic participants”, “analytic partners” and “analytic interlocutors” will be used interchangeably.

\(^2\) By “energetic” I refer to states of bodily energy, mostly to the processes of build-up and release of tension. These are closely related to the affects.
Therefore, it is argued that an elucidation of the nonlinguistic sonic dimension in the analytic encounter will provide important insights that are relevant to the psychoanalytic situation.

The research is based on the observation that, despite of the centrality of both listening and the voice in psychoanalytic therapy, their role and meaning have been largely overlooked by psychoanalytic theory.\(^3\) The psychoanalytic neglect of the vocal-auditory dimension is apparent not only in relation to the clinical situation but also regarding the field of music and the role of voice and audition in human development. An investigation of the literature leads to the assumption that the marginalization of the sonic realm in psychoanalysis is related to three major factors: First, the centrality of language in the psychoanalytic investigation of psychic life and human consciousness. Within this linguistic emphasis, psychoanalysis implicitly viewed the voice as a mere tool for the production of speech, failing to address its meaning and role as a phenomenon in its own right. In this capacity, the voice as sonorous object was noticed only in outstanding cases, when it exceeded its “normal” linguistic role. As elucidated by the Lacanian philosopher Mladen Dolar (2006), this position accords with the phonologic tradition that related language to vocal sound, however paradoxically effacing its effects of presence as soon as linguistic meaning is achieved. The obliteration of the voice was further reinforced by the linguistic turn in twentieth century Western thought, which questioned the epistemological nature of language and problematized the role of speech in

\(^3\) In this study I will refer to methods of therapy based on psychoanalytic theory. In addition to psychoanalysis, I will refer mainly to psychoanalytic psychotherapy.
interpersonal and cultural contexts (Dolar, ibid.; Kimbrough, 2002). As both these authors argue, Derrida's critique of the voice as a faulty element in the phonologic tradition was pivotal in deconstructive and poststructural theory. Both theories challenged the role of speech and the traditional sense of agency afforded the voice and the speaking subject (famously by Rousseau, [1781/2008] whose text is the central reference in Derrida's foundational text *Of Grammatology* [1998]).

The second factor in the marginalization of the sonic realm is the visualistic bias that characterizes Western culture. Vision was privileged as early as the ancient Greeks and became related to knowledge and enlightenment. Aristotle stated that “above all we value sight ... because sight is the principal source of knowledge and reveals many differences between one object and another.” (cited in Ihde, 2007, p. 7). Phenomenologist Don Ihde points out that these words stress the association between knowledge, vision and differences and distinctions between objects. “The preference for vision is tied to a metaphysics of objects. [In the Greek conception], vision already is on the way to being the ‘objective’ sense.” (ibid.). As we shall see, the sonic realm is fundamentally at variance with the visual orientation of distinctness and difference. As argued by several thinkers, this powerful visuo-linguistic tradition facilitated the marginalization of other modes of human experience, especially the auditory. These were consequently removed from cultural discourse (De Certau, 1984; Ihde, 2007; Corradi- Fiumara ,1990; Frie, 2003; McLuhan, 1962; Ong, 1982/2002). As a result, our current account of human experience is constructed - and reified - through visuo-linguistic terms. The visuo-

---

4 The privileging of vision has evolutionary reasons as well, as explained by, for example Straus (1966).
linguistic slant did not skip psychoanalysis. With its emphasis on language (the unfolding of free association, the discursive contents of narration, conversation and interpretation) and vision (e.g. the extensive use of spatial metaphors and the emphasis on the investigation of dreams), psychoanalysis has unwittingly affirmed this bias, neglecting to account for the centrality and significance of voice and listening upon which the clinical method is built.

This, however, is not the whole picture. Over the course of the 20th century, developments in psychoanalysis have problematized most of its central tenets, and in particular the centrality and role of language. Among these developments was the application of psychoanalytic treatment to nonverbal populations (such as children, autistic, borderline and psychotic patients). This inevitably directed the focus of attention to the nonverbal aspects of communication, raising the need to better understand and conceptualize this facet of human experience. The development of new, relationally oriented theories such as Object-Relations, Interpersonalism, Self-Psychology, the British Independents, and American Relational and Intersubjective theories decentered the importance of verbal narration and emphasized the actual therapeutic relation as the main locus of inquiry and lever of change, thus positioning the analytic dyad, rather than the individual patient, as the field of inquiry. In addition, the interdisciplinary dialogue increasingly conducted in the last decades between psychoanalysis and various

---

5 Throughout this study I use the term “relational” as an umbrella term that encompasses several psychoanalytic schools, such as Object Relations, Self Psychology, Interpersonal, Intersubjective and Relational views. The common denominator of these theoretical orientations is that they see object relations as primary rather than derived from intrapsychic drives (see Taub, 2009).

6 The scope of the present research does not allow an in-depth survey of these changes and their impact on classic psychoanalytic tenets. These are extensively accounted for in the literature (for example in Mitchell, 1993; Stolorow, 1997, 2003).
disciplines across the humanities and sciences (philosophy, neuroscience and especially developmental research), has voiced a strong critique regarding the linguistic and positivistic foundations of classical psychoanalysis (see for example Frie, 1999, 2003; Gendlin, 2003; Jacobs, 2001; Mitchell, 1993; Stern, 1989; Stolorow, 1997, 2003). Among various claims, these authors contend that the meaning of the body and of embodied experience⁷, especially in the clinical situation, has been poorly accounted for in the psychoanalytic project. They indicate the dire need for an inclusion of these factors in the psychoanalytic account, as well as the revisiting of central analytic concepts such as subjectivity, the self and the unconscious, which, as they argue, have been reified through spatial and mechanistic metaphors. As I will discuss later on in this chapter, modern developmental research has contributed much to the understanding of the nonverbal dimension of dyadic communication, and several ideas emerging from this scholarship have been successfully integrated in current psychoanalytic discourse. All in all, while these changes have enhanced current psychoanalytic minding and recognition of the experiential nonverbal realm, it is observed that the main conceptual innovations relate to the visual (facial and gestural) and bodily-symptomatic level, while the issue of sound (listening and the voice), although recognized as significant, has nevertheless remained insufficiently elucidated (however, one finds exceptions in Knoblauch, 2000; Pally, 1996, 2001; Rizzuto, 1995, 2002; Rose, 2004; Stein, 1999; 2007; Stern, 1985).

This brings me to the third factor that contributed to the psychoanalytic neglect of the sonic realm – its fundamental structure. Being primarily temporal, experiential,

⁷ The notion of embodiment and lived-experience will be elucidated in chapter II.
embodied and nondiscursive, sound famously defies verbal articulation. Its specific structure places the sonic realm in direct conflict with values of clarity, distinctness, repeatability, and quantifiability that are central to the traditional Western scientific and philosophic value system. The difficulty in articulating and conceptualizing the sonic order is also renowned in music theory, and the question of musical meaning remains one of the central challenges in this field.

The lack of conceptualization of both experience and music in psychoanalysis suggests a significant link between the two. And indeed, the structural affinity between the sonic order and lived-experience has been repeatedly indicated (for example by the aestheticians Suzanne Langer [1953, 1970] and Victor Zuckerkandl [1957]): both are temporal, nondiscursive, prereflective, embodied, happening in the now, and both resist verbalization. The irreducible gap between these phenomena and their conceptualization is another important trait they share. As phenomenologist and psychologist Eugene Gendlin (2003) indicates, “it is impossible to grasp the relationship between experience and concepts by concepts alone. Experiencing cannot be copied, captured, or represented. Concepts can only point toward our experience. Thus, we need to form concepts of a special kind that incorporate experiencing itself. ... Since experiencing cannot be represented, the concepts can only indicate various kinds of relations between experiencing and conceptual patterns” (Gendlin, 2003, pp. 100-101). As I will demonstrate in chapter II, however, this apparent ineffability is not absolute, but is rather a part of the traditional visuo-linguistic account, that tended, in self-affirmation, to mystify those facets of human experience (like music) that did not fit its percepts. These properties of lived-experience as conceptualized by phenomenology and the scholarship on subjectivity will be further elucidated in chapter II.
IV, metaphors or, in the present case, music metaphors, are an instance of “concepts of a special kind” that point out a certain experiential quality in the analytic encounter. But the irreducible gap between experience and language must not deter theorists from expanding the former’s understanding and conceptualization. This is of paramount importance in the psychoanalytic field, since it is the analytic experience that is the central force in effecting psychic change. It is therefore believed that a better conceptualization of the sonic dimension in psychoanalysis will contribute important insights to the understanding of qualities of general experience.

At this point, I wish to indicate that this dissertation emerges out of my personal experience of many years as a musician, music therapist and psychoanalytic psychotherapist. In these various professional capacities, the question of the relationship between sonic and musical experience and their verbal conceptualization was persistently - and increasingly - troubling. The various elucidations of this conundrum offered by psychoanalysis and music therapy had always struck me as insufficient and moreover - reductive, since they are for the most part embedded in Western culture’s traditional binary of body and mind. In this dichotomy either one of the poles has a tendency to dominate the other: While psychoanalysis privileges verbal insight, music therapy claims primacy to sonic nonverbal experience. The present dissertation represents my attempt to reconcile these extremes, providing a non-dichotomous, multidimensional model of the sonic dimension and its various manifestations, where experience and its conceptualization interact in a dynamic, reciprocal manner. In this model sound, the
voice and listening are repositioned in a fluid continuum, producing forms and manifestations of various representational and symbolic value (thus, although the study does not address music therapy in a direct manner, I believe it can be contributive to this discipline, which has been haunted by this binary split since its inception. See also fn. 19).

In light of the above, the main purpose of this study is to relocate the nonlinguistic sonic dimension in a more central position in the psychoanalytic clinical scenario and conceptualize its role and meaning. This entails an elucidation of the operation of voice and listening, both in terms of their phenomenology and in their configurations and mechanisms of signification. In the next section, before presenting the theoretical scheme of this study and its methods, I shall describe and qualify the nature of the nonlinguistic sonic dimension.

I.2 Description and Characteristics of the Nonlinguistic Sonic Dimension

The following is a basic phenomenological delineation of the elements comprising the nonlinguistic sonic dimension as viewed in this research. It includes:

a. The voices of patient and therapist in the course of conversation. This includes all voice components (pitch, volume, timbre and duration) and

---

10 Since my basic position is that phenomena are intrinsically related to the structure of their perception, in this study, the term "sonic" implicitly includes audition as its inherent perceptual counterpart (however, listening and audition will be addressed separately in chapter IV). The term "sonic" is often used interchangeably with "vocal". In addition, the term "sound" includes in the present account all of sound's features: timbre, pitch, volume and duration, and used in relation to both vocal and instrumental sounds of defined pitch (musical) and undefined pitch (such as noise and speech).
nonverbal sounds occurring in the conversation such as sighing, crying, laughing, coughing, and the like.

b. Elements and factors inherent in vocalization: each individual’s unique organization and articulation of sound through time (rhythm) as well as breathing and various manifestations of silence.

c. Vocal-temporal aspects of the dialogue such as vocal correspondence, matching, synchronization and modes of turn taking.

d. The process of listening and interpretation.

It is emphasized that the nonlinguistic sonic dimension is conceived primarily as a global entity emerging from the dyadic interaction. Thus, certain categories I discuss in coming chapters, like vocal ambience and vocal duet, are direct products of this interaction. But even those sonic phenomena produced individually by each of the individual participants, such as vocal gesture and vocal persona, are here viewed as essentially dyadic in that they are motivated by the dialogical matrix. This view accords to a basic conception of the psychoanalytic situation as a whole – the “total clinical situation” (Joseph, 1985) - as the main object of investigation. This conception is consistent with the contemporary view of psychoanalysis as a “two-person psychology” operating as a nonlinear, dynamic dyadic system. According to this view, the human mind is conceived as fundamentally dialogic (Beebe and Lachmann, 2003a, Beebe, Knoblauch, Rustin & Sorter, 2003b; Mitchell, 1993, Mitchell and Aron, 1999; Stolorow, 1997). The meaning of this is that, in contrast to the traditional psychoanalytic focus on the patient's voice, and the ascription of listening mainly to the analyst, patient and therapist are here viewed as co-
participants in a mutual and reciprocal vocal-auditory activity in their creation of the sonic dimension (this does not imply that their activity is identical, nor is it symmetrical. On the contrary, I shall discuss in the course of this study how their different roles in this situation determine in many ways their vocal-auditory styles and behaviors).

Three central features characterize the nonlinguistic sonic dimension:\(^{11}\):

1. **Sound's formal structure.** Sound's unique characteristics differentiate it from material objects: it unfolds in time, it is continuous and dynamic. These make sound a special, *processual*, type of sign. In addition, sound lacks distinct boundaries: it spreads in space and intermingles with other sounds. These are the very properties that make the flow of sound difficult to segment and articulate. As we shall see, the listening process – the perceptual counterpart of sound – is determined by these qualities.

2. **Sound's meaning.** Sound is characterized by being abstract, nondiscursive and basically nonrepresentational, making its effects nearly impossible to paraphrase or otherwise verbalize\(^{12}\). These characteristics were famously debated in music theory, in an attempt to discern the meaning of music. The question of sound's signification constitutes one of the central challenges to be met in the present study.

---

\(^{11}\) These features are presented here in general form and will be addressed in greater detail in the next chapters.

\(^{12}\) Naturally, certain aspects of sonic phenomena can be represented, for example by imitation (as in the musical imitation of wind, water, horses galloping, etc.). However, this form of representation was demonstrated to be inessential to the signification of sound and music (Kivy, 1980; Langer, 1970; Zuckerkandl, 1957).
3. **Sound’s presentation.** Sound’s presentation and reception is fundamentally experiential, embodied and prereflective - in other words, its occurrence in present time is essential. These properties require the clarification of three important issues: a. the role of the body in acts of performance and perception; b. the nature of present-time consciousness and c. the interpersonal context in which sonic events unfold. As we have seen above, these are precisely the issues that psychoanalysis, due to its visuo-logocentric slant, has dealt with in an insufficient manner.

Before proceeding to illuminate the nature of the nonlinguistic sonic dimension, a caveat, concerning its relation to the linguistic and the paralinguistic aspect of the voice, is required. The global phenomenon of vocal sound in communication is considered in the present study to be a complex, multidimensional entity that functions in several signifying capacities, employing various - and different - sonic components and mechanisms that simultaneously generate various meanings. The *linguistic voice* employs phonemic components (such vocal components as pitch, stress, vowels and consonants) to produce phonemes. These compose, in turn, words and sentences. These sonic signs are unmotivated, i.e., they are arbitrary, conventional (agreed upon by the speakers of a given language), and hold no resemblance to their referents. The phonemic structure of words renders them discrete, facilitating segmentation, articulation and a more determinate ascription of meaning. The linguistic sonic level is characterized by distinct vocal articulation and a linear and sequential deployment. As demonstrated by de Saussure (1916/1983),

---

13 In this study I exclusively address sonic events that transpire between humans, as in live musical performances and conversations. The subject of sonic experience that comes about through technological means is a separate subject that cannot be addressed presently.
words are constructed on a sonic differential-binary principle that facilitates their
distinction and the denotation of stable and conventional meanings (that are also,
connotatively, relatively stable). As phrased succinctly by Dolar: “the [linguistic]
signifier possesses a logic, it can be dissected, it can be pinned down and fixed - fixed in view of its repetition, for every signifier is a signifier by virtue of being repeatable, in view of its own iterability.” (Dolar, 2006, p. 16-17). Obviously, the perceptual correlate of linguistic sound - linguistic listening - is also logical, cognitive, symbolic and reflective.

The paralinguistic voice is defined as the set of non-phonemic properties directly related to the performance of speech, such as speaking tempo, loudness, stress, vocal pitch and intonational contours. The paralinguistic level of speech involves conventional or dialectic manners of verbal performance within a given language. At times, these are also used more idiosyncratically, or idiolectically, to communicate emotional attitudes and various, more covert shades of meaning. At any rate, in the paralinguistic order, the voice is subjugated to the verbal realm, placing it, as it were, in the exclusive service of verbal discourse, as its designation readily suggests.  

As Dolar’s rather sardonic account goes, in the phonologic operation, paralinguistics is the “surplus that is necessarily added to the purely phonemic distinctive features –

14 In the history of relations between voice and word, the designation of non linguistic vocal expression as para-linguistic reflects the politics of language: the prefix ‘para’, appearing in loanwords from Greek, means: “at one side of, beside, side by side” (parabola; paragraph; parallel); “beyond, past, by” (paradox). Subsequently, it came to designate objects auxiliary to or derivative of those denoted by the base word (parody); or the abnormal or defective (paranoia), a sense now prevalent in modern scientific terms. It is also employed in the naming of occupational roles considered ancillary to roles of a higher status (paramedical).
the prosody, the intonation and the accent, the melody, the redundant elements ...

Bones, flesh, and blood of the voice were diluted without remainder into a web of structural traits, a checklist of presences and absences.” (Dolar, 2006, p. 19).

The present conception of the nonlinguistic sonic dimension differs from the concept of paralinguistics (though it does not seek to undermine it). It refers to a realm of signification that exceeds the linguistic: without contributing directly to the immediate verbal semantics, it generates another set of forms, configurations and meanings that are partially independent from the verbal flow. In addition, the forms and articulations of the nonlinguistic sonic dimension often do not correspond formally to the verbal level. For example, the phenomena of vocal timbre, vocal persona and vocal ambience (these will be elucidated in chapter III), are vocal configurations created by the interplay between the phonemic and non phonemic components of sound that significantly affect the communicational field and are, to a large extent, independent of the immediate semantics and structure of the verbal flow. Although the present conception differentiates between the paralinguistic and nonlinguistic aspects of the voice, it is noted that both aspects elide and interweave continuously, blending in various ways.

As pointed out above, the nonlinguistic sonic dimension is experienced in a fundamentally embodied, pre-reflective, participatory way. Characterized by being continual, non-discrete, non-propositional and non-discursive, this dimension is given phenomenally as a diffuse sonic flow whose configurations are indeterminate and multivalent, constantly in the making. Its meanings are highly unstable, largely context-dependent and established ad-hoc. The segmentation of this level, as well
as the ascription of meaning to its unfolding are mostly delegated to ongoing impressionistic, subjective interpretations that unfold in a sentient, embodied manner in the course of mutual listening, continually integrating meanings from other (linguistic and paralinguistic) levels of signification. The experience and interpretation of the nonlinguistic sonic dimension positions the listener at odds with the order of consensual, determinate meanings and introduces her to a realm of participation, interpenetration, impression, uncertainty and qualitative texture. The intricacy and tension in the dialogue between these various levels of meaning are addressed in chapter II through the dialogic theory of philosopher and literary critic Mikhail Bakhtin.

As described above, the nature of nonlinguistic sound raises essential questions about its signifying mechanisms and principles of segmentation and articulation, as well as the nature of the listening, perception and interpretation it engenders. A first step in such an elucidation is the inquiry into the psychoanalytic relation to sound and music and, more explicitly, sound in the context of human development and communication.

### I.3 Sound and Music in Psychoanalysis

It has been argued in psychoanalytic and related literature that classic psychoanalysis has dealt with the issue of sound and music in an insufficient manner and that the affective power of music has remained, from a psychoanalytic point of view, elusive and only partially explained (Langer, 1970; Noy, 1966, 1967a, 1967b,
Freud’s visual bias coupled with his alleged dislike of music has been referred to as a decisive factor in the seeming neglect of this subject, although his famous aversion to music has been debated lately (Cheschire, 1996; Diaz de Chumaceiro, 1990, 1993; Feder, 1998). As we have seen above, the reasons for this neglect seem to be related more to cultural and epistemological factors and to the inherent structure of music than to the personal inclinations of this or that researcher. Nass (ibid.) suggests that the lack of psychoanalytic writing on music may be ascribed to the nondiscursive nature of music as well as to the musical illiteracy of most psychoanalysts. Indeed, it is observed that the most significant texts on psychoanalysis and music were written by psychoanalysts who are also musicians or have a strong affinity to music (see for example Feder, 1990, 1993; Nass, 1971, 1975, 1984, 1989; Stein, 1999, 2004a, 2004b, 2007; Rose, 1993, 2004).

Although a considerable body of writing has accumulated over time on the subject of music, the general impression is that these writings constitute relatively isolated attempts at theorization and do not amount to a coherent psychoanalytic conception of the sonic-auditory sphere and its various capacities in human experience. Issues such as the meaning of sound in the musical context, its connection to affect, its representational mechanisms and the musical creative process have not been sufficiently or convincingly elucidated. A central problem is that for the most part, psychoanalysis addressed music in isolation from the auditory and vocal realm in human development and particularly separated this issue from the sonic realm in the clinical setting. It is quite remarkable that a therapeutic
procedure that uses voice and listening as its main vehicles of expression did not engage in a more thorough investigation of these issues, nor did it create meaningful links between these three areas (sound in development, music and the clinical setting). However, a chronological observation of the trajectory of psychoanalytic discourse on sound and music indicates that this traditional orientation is changing.

Classical psychoanalytic literature has dealt with music as an art form, focusing on subjects such as the auditory sphere, the meaning of musical pleasure, the creative drive of composers, analyses of musical compositions and psychoanalytic views of musician’s lives. As Nass (ibid.) and Noy (ibid.) note, the analytic concepts employed in this literature reflect the basic psychoanalytic approaches in the particular time of their writing. Earlier papers conform mostly to early Freudian or later ego-psychological orientation, emphasizing the adaptive-defensive functions of musical activity, viewing music as a sublimatory transformation of unconscious aggressive, erotic or regressive drives and conflicts (see for example Brody, 1943; Ferenczi, 1919; Kohut, 1950; Kohut and Levarie, 1957; Reik, 1946, 1953). An extensive review of this literature is found in Feder, et al. (1990, 1993), Noy (1966, 1967a, b, c, d) and Nass (1988). A critical review of this direction of inquiry can be found in Nass (1988), who indicates its reductionistic slant. This orientation is influential in early psychoanalytic theorization on the voice and the auditory sphere, as elucidated in chapter III.

With the advent of relational psychoanalytic theories, sound and music have come to be conceptualized through an interpersonal theoretical lens. Accordingly, sound has been argued to represent the maternal object or the object relation. In this
light, sound and music have been theorized as transitional phenomena (Winnicott, 1971), transformational objects (Bollas, 1987a, 1992), an audio-phonic skin (Anzieu, 1979, 1990), a type of psychic envelope (Lecourt, 1990) and a proto-representation of the object (Maiello, 1993, 1995). Although the scope does not permit the explication of these important concepts, it is stated that these theorizations are most valuable for two reasons: first, they position human sound and its later artistic and cultural manifestations in an interpersonal context. In this light, sound’s relational, addressive and communicational intent is highlighted, rather than its adaptive or intra-psychic symptomatic meanings. The latter, though valid, are not sufficient to explain sound’s meaning, as I have maintained above. Second, these new conceptualizations highlight the interpersonal, intersubjective dynamics underlying the creative act, and facilitate bridging between sound in human development and the creative impulse and their coalescing, later in development, into musical artistic activity (see especially Ehrenzweig, 1965; Milner, 1987; Nass, 1971, 1975, 1984, 1989; Rose, 1993, 2004; Winnicott, 1971). Nevertheless, the above conceptualizations are only partially relevant to the present study, since they are theoretical by nature and do not delve into the concrete unfolding of sound in the clinical situation.

In a secondary way, psychoanalysis has also dealt with the subject of sound and the voice within the analytic setting. Most of the articles relating to this context deal with music and its unusual appearance in the analytic setting, such as the patient’s associations to music, songs, or such events as humming, whistling or referring to music within the analytic setting (Bryce-Boyer, 1995; Mitrani, 1995; Greenson, 1954;
Hannett, 1964). Relatively little material seems to exist on the subject of the significance of the voice and vocal expression as sonorous events in the moment-to-moment unfolding of analytic communication, as well as their various derivatives such as silence, rhythm and the listening experience, and their contribution to the analytic process. As critiqued in chapter III, most articles addressing voice and listening in this capacity are linguistically biased and reductive in their approach. However, in the course of psychoanalytic evolution, substantial links between sound and its components in human experience and the clinical setting have been recognized by an increasing number of analysts (significant examples are: Anzieu, 1979; Baranger, 1993; Bion, 1963; Blos, 1972; Bollas, 1987b; Coltart, 1993a, 1993b; Faber, 1988; Greenson, 1961; Hadda, 1991; Joseph, 1989; Khan, 1974a; Killingmo, 1990; Klein, 1961; Knoblauch, 2000; Lecourt, 1990; Landau, 1996; Maiello, 1995, 2000; Nass, 1971, 1989; Ogden, 1994a; Paul, 1989; Priel, 1997, 2003; Rizzuto, 1995, 2002; Rose, 2004; Sabbadini, 1992; Schwaber, 1983; Spence, 1984; Stein, 1999, 2007; Tustin, 1986; Winnicott, 1989, 1971). Several of these writers (notably Feder, et al., 1990, 1993; Knoblauch, 2000; Lecourt, 1990; Maiello, 1995, 2000; Nass, 1971; Rose, 2004; Stein, 2007) have made more explicit connections between clinical happenings and music. In addition, one of the significant phenomena discerned in the literature, is the increasing appearance of musical metaphors in the description of psychoanalytic experience. This issue is addressed in chapter IV.

Recent psychoanalytic conceptualizations of music and sound are more interdisciplinary in nature and are provided by various authors, such as Knoblauch, 2000; Lachmann, 1999; Rose, 1993, 2004; Stein, 1999; Stern, 1998; and Storr, 1992.
These support more explicitly integrated perspectives, combining music theory, infant research, aesthetics and neuroscience. This interdisciplinary impetus has brought about a more consistent theorization of sound and music in that it amalgamates developmental roots and aesthetic considerations, and most importantly, has created substantial theoretical links between the various disciplines that deal significantly with sound in human experience.

An additional source of psychoanalytic information that, albeit indirectly, informs the investigation of sound in the present study is the growing literature on the nonverbal aspect of human life. Also referred to as the implicit, pre-symbolic, sub-symbolic, somatic, out-of-awareness or un-symbolized aspect of psychic life, this issue has come to the psychoanalytic limelight in the past four decades (see e.g. Alvarez, 1992; Anthi, 1986; Beebe and Lachmann, 2002; Bollas, 1987a; Bucci, 1997; Coltart, 1993a; Fosshage, 2004; Joseph, 1985; Khan, 1974b; Mancia, 2003a, 2003b; Mitchell, 1993; Mitrani, 1995; Ogden, 1989, 1990; Pally, 1996, 2001; Ringstrom, 2001; Stern, 1985, 1998, 2004; Winnicott, 1984b, 1984c). This direction of analytic inquiry is an expansion of the more central psychoanalytic focus on the investigation of the unconscious as revealed through verbal narration. Through this, various conceptualizations of the nonverbal domain have been made. Some of them are concerned with the symptomatology of somatic, developmental and characterological disorders (Anzieu, 1990; Bick, 1988; Coltart, 1993, 1996; Khan, 1998).

---

15 Though the nonverbal aspect of experience, especially as it is expressed through discrete physical symptoms, has been one of the original preoccupations of psychoanalysis since its very inception, as early as Freud’s “Studies on Hysteria” (Breuer and Freud, 1893-95), it seems that psychoanalysis was less attentive to, or perhaps lacked the tools to analyze, the more fluid and continual somatic, gestural, vocal and rhythmic aspects of the analytic process that occur in the moment-to-moment of the analytic session. These are the very aspects that have an apparent affinity to music.
In this light, phenomena related to breathing, eating, digestion, the skin, the voice, bodily rhythms and gestures and silence and sleep within the analytic setting are understood to be a form of non-symbolic or pre-symbolic representation of unconscious mental-affective themes, specifically reflecting early mother-child interactions and/or unmentalized trauma. It is noted that these conceptualizations consistently address circumscribed bodily symptoms and other visible manifestations of pathology, whereas the vocal realm is again, largely left out of this growing discourse.

However, the vocal auditory realm seems to be somewhat restored by the novel focus on the nonverbal as manifested in the real or actual therapeutic relationship as a somatic-affective lived experience. This literature investigates the more fluid aspects of the “moving along process” of therapy (Stern et al., 1998, Stern, 2004) and the unfolding quality of the “going-on-being” (Winnicott, 1984d) of the therapeutic meeting. Rather than on focusing on discrete physical symptoms, the emphasis is placed on the interactional flow: on here-and-now reciprocal processes of attention, accommodation, matching and response and on the analyst’s countertransference as a felt experience (the analyst’s somatic state, voice, etc.). Facial and gestural, vocal and rhythmic components are considered important factors in the evolution of the relationship. Furthermore, introjected relational patterns (object-relations) are believed to be implicitly enacted through vocal-temporal elements (by e.g., Alvarez, 1992; Balint, 1968; Beebe and Lachmann, 2002, 2003a; Beebe et al., 2003a, 2003b, 2004; Bollas, 1987b; Ehrenberg, 1992;
Knoblauch, 2000; Maiello, 1995; Ogden, 1994, 1994a; Rose, 2004; Ringstrom, 2001; Stern, 1985, 1998, 2004, Stein, 2007). This direction of inquiry is closely related to contemporary infant research, which I address in the following section.

I.4 The Justification for an Interdisciplinary Research

Since the psychoanalytic elucidation of sound has proved as only partially relevant to the present study, the scope is broadened to include the scholarship from developmental research and music theory. These areas of study have made significant explorations regarding sound and its affective, expressive and representational features but have done so in a relatively isolated manner, resulting in fragmentary viewpoints on similar subject matter.

1.4.1 The contribution of Developmental Research

In the last four decades, the study of development has investigated the processes of the establishment of a sense of self as emerging from a dyadic, interpersonal, intersubjective matrix. The scholarship in question is based on non-linear dynamic systems theory and informed by cognitive science and neuroscience (Beebe et al., 1988, 1997; Bucci, 1985; Fosshage, 2004; Malloch and Trevarthen, 2009; Pally, 1996, 2001; Rose, 2004; Schore, 1997; Stern, 1975, 1985, 1987; Stolorow, 1997; Thelen and Smith, 1994; Trevarthen, 1998, 2001). Rather than centering on the individual mind and its intra-psychic psychological processes, the focus here is on the observed
Human symbolic capacities are claimed by these researchers to be essentially dialogic\textsuperscript{16}, contextual and intersubjective. Part of this inquiry highlights nonverbal, nonrepresentational and intermediate representational states, in an attempt to identify the mental-affective mechanisms and processes employed to transform nonverbal experience into more complex symbolic structures. As we have seen above, these findings are increasingly linked to adult psychoanalytic treatment. It is contended that the nonverbal, implicit principles of self- and self-other organization and regulation characterizing the primary dyad continue to exist and activate adult partners, mostly in an out-of-awareness mode. This implicit level of interaction is argued to be more important than has been hitherto recognized (Beebe and Lachmann, 2002; Beebe et al. 2003a, 2003b, 2004; Lachmann and Beebe, 1996; Morgan, 1997, Priel, 1997; Stern, 1985, 2004). And indeed, it is evident that the body of knowledge emerging from developmental research is being increasingly integrated into current psychoanalytic theory, particularly in those strands that focus on the present analytic interaction and relation as a lever of change (such as the relational and intersubjective psychoanalytic movements in North America).

The importance of the above contributions to our present purposes is that it has been consistently demonstrated by this body of research that vocal-temporal and auditory components are central in the establishment of an early sense of self and self-and other within a meaningful relationship (Malloch, 1999; Malloch and

\textsuperscript{16} Although this line of inquiry does not delve into object relations as they are constructed in the unconscious, it nevertheless creates valuable bridges between the “observed baby” and the “psychoanalytically constructed baby” (Stern, 1985) as viewed through the object-relation prism, particularly in the Neo-Kleinian orientation.

\textsuperscript{17} In the present study, the use of the term “dialogic” is related to Bakhtin’s theory of dialogue, to be explained below.
Trevarthen, 2009; Robb, 1999; Stern, 1985, 1996, 2004; Trehub, 1990; Trevarthen, 1999-2000, 2002). Through the microanalyses of interaction of primary dyads these developmentalists have demonstrated the central role of musical elements in the creation of primary subjectivity and intersubjectivity. In particular, Malloch and Trevarthen (2009) have emphasized the musicality of the dyadic interaction of mother and infant and its importance in establishing a sense of human companionship and cultural belonging. Those writers who have endeavored to apply these findings and ideas in the adult analytic settings (in particular Knoblauch, 2000; and also Beebe and Lachmann, 2002; Beebe et al. 2003a, 2003b, 2004; Stern, 2004) have proceeded to examine interactional sequences in the present moment of the therapy, focusing on the vocal- temporal and dynamic contours of the dialogue. Through this outlook, patient and therapist are viewed as two semi-open systems that are in a state of continuous motion, mutual regulation and correspondence, resembling in many ways a musical dynamics, whereby musicians interact in a continual process of dialogue, reciprocation, reorganization and mutual reframing.

Currently, the psychoanalytic status of these nonverbal aspects appears to be in a process of evaluation and conceptualization, as they do not conform in an obvious way to the classic concept of the repressed unconscious (Freud, 1915). Several alternative formulations have been offered in this context, e.g. Bucci, 1985, 1997; Fosshage, 2004; Knoblauch, 2000; Mancia, 2003a; Mitrani, 1995; Pally, 2001; Tustin, 1990; Ogden, 1989; Stern, 1998, 2004. These formulations are strongly influenced by the different psychoanalytic orientations from which they originate.
1.4.2 The Contribution of Music Theory

The second discipline that is included in this study as significantly dealing with sound is the field of music theory. Music aesthetics and semiotics are considered in this research as key theories in the elucidation of the nonlinguistic sonic realm, that may provide insights regarding its principles and mechanisms of signification. The aesthetic-semiotic musical discourse is presumed to be relevant to the proposed study for several reasons: first the issue of sound as signifier and its organization into a form of communication has been central in music theory (Kramer, 2002; Langer, 1953, 1970; Zuckerkandl, 1957). This has potential importance for a psychoanalytic understanding of sound, especially regarding those features which psychoanalysis seems to have had most difficulty with, specifically music’s non-propositional, yet symbolic and narrative-like nature (see for example Cone, 1974; Cooke, 1959; Cumming, 2000; Dahlhaus, 1982; Epstein, 1993, 1995; HaCohen, 1999, 2003; Kerman, 1982; Kivy, 1980; Lidov, 2005; Meyer, 1956; Nattiez, 1990; Schwartz, 1997; Storr, 1992; Treitler, 1993; Zuckerkandl, 1957). In addition, concepts coming from music cultural studies (e.g. Auslander, 2006; Blacking, 1976; Echard, 2000, 2005; Keil and Feld, 1994; Shepherd and Wicke, 1997) and contemporary music semiotics (Cumming, 2000; Hatten, 1994; Lidov, 2005) significantly illuminate the relation between musical meaning and embodied experience. This link seems pertinent to the elucidation of nonlinguistic sound in the clinical situation, whose
primary signification is here believed to be fundamentally related to the embodied experience of the clinical situation. ¹⁸

Second, the aesthetic-semiotic discourse on the relations between music and affect (for example Cumming, 1996, 2000; HaCohen, 1999; 2001a, 2001b; Kivy, 1980, 1993; Langer, 1953, 1970; Monelle, 1992, 2000) can contribute to the understanding of these relations within the psychoanalytic context, especially considering that psychoanalysis has not yet produced a sufficient explanation as to the affective impact of sound.

A third underlying notion justifying the inclusion of music theory in this study, is the observation that the nonlinguistic sonic dimension of dialogue has a formal affinity to music. This affinity has been addressed in recent developmental researches and musical discourse, linking conversational and musical forms (Gratier and Apter-Danon, 2009; Lidov, 2005; Malloch and Trevarthen, 2009; Monson, 1996; Stern, 1996; Trevarthen and Gratier, 2008). For example, it seems possible to discern, in the course of a conversation, the emergence of sonic forms, themes and principles of organization that are suggestive of similar phenomena found in music – such as the principle of correspondence, and that of repetition and change (see chapter III). Phenomena such as vast variations in breathing, silence and vocal timbre, the creation of rhythmic and vocal gestures and characteristic forms of dialogue and reciprocation are evocative of musical forms such as theme and variations, fugal structures and responsorial chant, among others. The juxtaposition of music theory

¹⁸ I shall address specifically the strand of music-semiotics that is based on the ideas of the philosopher Charles Sanders Peirce, as elucidated below.
and psychoanalytic theory thus provides new ways to conceptualize and articulate various phenomena that unfold in the nonlinguistic sonic dimension.

Finally, the common roots of music and conversation are considered as validating the inclusion of music theory as a significant contributing field to the present study. In both verbal language and music, sound and time are the central elements interwoven in the transmission of the message. Both media employ the elements of sound intensity, duration, timbre and pitch in creating and structuring their particular texts. Another common feature is the activity of interpretive listening. For both verbal and musical texts, listening entails - beyond a mere decoding of a message - an active, ongoing construction and deconstruction of what is heard, and is inevitably colored by the listener’s subjectivity. Both language and music are here viewed as entities of a contextual and dialogic nature, containing within their very structure the presence of a listener - an interpreting subject, activated by them and activating them in turn.19

19 The field of music therapy is relevant to this particular point (and generally to subject of the present study), as it is positioned at an inter-disciplinary junction where music, developmental theories and psychodynamic theory converge. Music therapy literature discusses the function of sonic and musical expression in nonverbal communication and its potential as an alternative communicative medium with a wide range of patients that includes individuals with severe impairment in their motor, neurological, mental, emotional, and communicative abilities. The literature elucidates the therapeutic capabilities of sound, voice and music and their role in structuring a therapeutic relationship. In most clinical approaches, the use of musical free-improvisation by therapist and client is a prevalent method, and is employed as a diagnostic tool as well as a means for expression and communication. Ultimately, music therapy was not included in the research, for the following reasons: a. The theoretical scope of the study – psychoanalysis, phenomenology, music aesthetics and semiotics and infant research - proved to be too large and complex to allow the inclusion of yet another discipline; b. In addition, music therapy does not have its own meta-theory, and is itself embedded in other theories – most of which are mentioned above. Its inclusion would thus entail unnecessary repetition of theoretical material; c. Initially, music therapy’s most relevant contribution to the present study was considered to be its clinical aspects (such as clinical reports and various musical techniques for treatment). However, these concern therapy with nonverbal patients, a large part of whom suffer from severe mental and psychological impairment and pathology, while the present research focuses on psychoanalysis with verbal, non impaired patients. Since my central focus is investigating the manner in which the nonlinguistic sonic dimension interweaves with the linguistic
In order to conduct the proposed interdisciplinary inquiry, a particular conceptual context is created, integrating three main theoretical perspectives through which sonic phenomena in speech can be explored. This conceptual context is attuned to the specific, unique nature of sound, and facilitates its consideration as a fundamentally experiential, embodied, dialogic type of sign, bringing to the fore the shared features of verbal language and music, their common developmental sources and dialogic relations as well as the distinct functions of sound in each media. The conceptual matrix I propose is formed out of a conviction that any conceptualization of the nonlinguistic sonic dimension must have an inherent pertinence to sound's characteristic fluid and evanescent nature. In this sense, the conceptual system I suggest acts more like a “facilitating environment” (Winnicott, 1971) in which ideas on sonic phenomena can thrive and expand, than a theoretical instrument used to decode or dissect such phenomena. The environment I try to create is in itself suggestive and dialogical. Because of its interdisciplinary nature, some of the tenets on which one individual theory is built, may not be wholly compatible with tenets of another. But these contradictions and dissonances do not obstruct the dialogic process between other parts of these theories; on the contrary, they create a dialogic matrix that facilitates certain sonic phenomena to resonate better; in turn, the level, the contribution of music therapy was revealed to be less pertinent to the subject of the research than initially considered. The reader is referred to representative literature (Alvin, 1977; Amir, 1990, 1999; Ansdel, 1995; Austin, 1999; Bonny and Savary, 1990; Bruscia, 1987, 1989, 1991; De Backer, 2004; Elefant, 2002; Nordoff and Robbins, 1977; Odell-Miller, 2001; Pavlicevic, 1990, 1997; Sekeles, 1995, 1996, 1996a; Smeijsters, 2003; Streeter, 1999a, 1999b, Wigram and De Backer, 1999; Wigram et al. 1995, 2002; Frank-Schwebel, 1996, 2001, 2002, 2002a; Heal and Wigram, 1993, Hughes, 1995; Metzner, 1999; Nordoff and Robbins, 1977; Priestly, 1975; Robarts, 1996; Sekeles, 1995, 1996, 1996a 1996; Sutton, 2002a, 2002b; Wigram and De Backer, 1999).
this newly constituted conceptual matrix becomes gradually more vital and attuned to the phenomena in question, through its ongoing dialogue with the conceptualizations it helped create.

The suggested conceptual environment is receptive to all manner of sonic phenomena: from global ambience and character to local textures and gestures, fleeting sonorities and dynamic motion that come into being only to fade away, leaving resonances that can barely be heard and even then only inwardly and imaginatively. This environment is one that must be able to accept and handle such evanescent experiences gently, without recourse to “violence of interpretation” (Aulagnier, 1975/2001) and without “robbing” them of their essence (as phrased by Levinas, cited in Kenaan, 2009).

From both a philosophical and a methodological point of view, I address the sonic dimension with three objectives in mind: first, to develop an approach that is able to attend to, but also maintain and nurture the inherent global, processual nature of the nonlinguistic sonic order; second – to be able, while preserving this approach, to address more local levels of sonic detail, leading to an articulation of sound’s expressive and rhetoric operation (i.e. identify categories, structures, particulars of presentation and mechanisms of signification); finally, because sound is viewed as an essentially dialogical entity, my conceptual framework engages the interlocutors as expressing and negotiating various forces and intentions. My focus is on the relational quality and the intentionality and directional impulse within relationships.
This theoretical strategy endeavors to defer traditional linguistic attitudes and to avoid the need to break down sonic phenomena through binary-linguistic or positivistic methods. Well-known binaries usually addressed as dichotomous and antithetical (for example: text and listener, words and music, the verbal and non-verbal domains, primary and secondary processes) are positioned on one conceptual continuum, and their various manifestations are viewed as blends and configurations of different modes of representation (see Noy, 1999). In effect, my approach contains an inherent critique of those attempts to understand the process of sonic signification exclusively through formal analyses and other visuo-linguistic methods. These are believed to be insufficient and at times inappropriate (leading to grave consequences regarding sound's fate\textsuperscript{20}) to explicating the effect of sound with its temporal, global, qualitative and indeterminate nature and its expressive function in subjective and intersubjective emotional communication. In accordance, I shall use the following ideas in a loosely coordinated manner, borrowing freely and making adaptations as needed.

The first theoretical perspective and decisive philosophical point of reference is phenomenology, the philosophical theory and method conceived by the German philosopher Edmund Husserl during the late 19th and early 20th centuries. Phenomenology, in Husserl's conception, is primarily concerned with the systematic reflection on and analysis of the structures of consciousness and the phenomena which appear in acts of consciousness. Such reflection takes place from a generalized first-person viewpoint, studying phenomena not as they appear to “my”

\textsuperscript{20}In his consideration of the voice within structural linguistics, Dolar wittily makes a pun with the Greek words phone (voice) and murder (phonos) remarking that phonology sought to kill the voice: “Phonology stabs the voice with the signifying dagger” (Dolar, 2006, p. 19).
consciousness, but to any consciousness. In this study, phenomenology will be used as a fundamental stance towards sonic phenomena in the psychoanalytic situation. At times, however, it will also be used as method to explore the structures of sonic phenomena and sonic consciousness, or the manifestations of the nonlinguistic sonic dimension in the analytic conversation in the listener’s experience and the meanings it generates. My approach to sound in speech and especially to the dialogic aspect of the encounter is particularly based on the thinking of Merleau-Ponty (1962/2008), whose philosophy of perception continues and expands the ideas of Husserl. Several concepts developed by Merleau-Ponty (especially the concepts of being-in-the world, lived-experience, the body-subject and the embodied-self – to be explained in the next chapter) will provide the basis for my understanding the sonic dimension of the psychoanalytic encounter. In addition, I will refer to contemporary scholarship on consciousness, subjectivity and intersubjectivity in the fields of psychoanalysis, philosophy and cognitive science (for example Frie, 2003; Gallagher and Zahavi, 2008; Zahavi, 2000, 2004, 2005) that expands Husserl and Merleau-Ponty’s ideas. These contemporary authors further elucidate the phenomenological concepts of the embodied self and embodied experience, focusing especially on the issue of prereflective consciousness and the embodied mind. The nonlinguistic sonic dimension will be addressed primarily within this conceptual framework.21

In contrast to the phenomenological approach that commences its investigation from within the experiencing subject, the second perspective I employ discerns the

21 The differences between this conception of consciousness and unconsciousness and the classical concept of the unconscious as it is conceived in psychoanalysis will be explained in chapter II.
operation of sound from without, by formulating more objective, relatively distinct sonic categories and articulations. For this purpose, and in light of the idea that the nonlinguistic sonic functions as a semiosis, I turn to music semiotics as a second point of reference. The latter is introduced into the psychoanalytic arena in an attempt to elucidate processes of signification of the nonlinguistic sonic dimension in speech that differ in structure from that of the linguistic, as we shall see.

The foundation for the third theoretical perspective is Mikhail Bakhtin's theory of dialogism. Although Bakhtin's ideas emerge within the framework of literary theory, his work expanded well beyond that range, forming a modern epistemology that addresses human behavior in society and culture through the use of language. Bakhtin's relentless endeavor to address the human subject's unique individuality and agency is cardinal to the present study. In this, Bakhtin stands in opposition to postmodern trends which deconstruct the subject and annihilate her unique, unrepeateable individuality. Bakhtin offers concepts that posit a continual open, non dialectic and non unitary juxtaposition of the individual and the social forces operating upon her. The Bakhtinian dialogic matrix, I contend, is a key to the understanding of nonlinguistic sound in analytic communication. Within this conception, the sonic dimension is addressed as being fundamentally motivated by the dialogic context in which it unfolds; a context that is infinitely layered and includes not only the concrete interlocutors, but additional (personal, historical, social and cultural) forms of dialogue that reverberate in the actual encounter.

Nevertheless, as I will demonstrate, the distinct disciplines of phenomenology and semiotics are connected in several ways and complement each other throughout this suggested approach.
Bakhtin’s dialogic conception and its correlating notion of addressivity, inform my understanding of sound’s impact through performance in speech, as well as the agentic power of the voice in an intersubjective sonic milieu. These concepts will be explained in depth in the next chapter.

Finally, it should be noted that this conceptual context reflects and enhances an ongoing discourse that has been unfolding in the past decades in the three main bodies of knowledge that compose this research: psychoanalysis, music theory and infant research. Phenomenological philosophy, music semiotics, and Bakhtin’s dialogism are engaged, though each in different ways, in these disciplines in the investigation of human experience and communication and have increasingly appeared as a meaningful conceptual base in these fields. (Examples of such interdisciplinary discourse is found in Beebe and Lachmann, 2002; Canestri, 2000; Corradi-Fiumara, 1990; Harris and Aron, 1997; Lidov, 1999; 2005; Priel, 1999, 2003; Rizzuto, 1995; Shapiro, 2000; Shepherd and Wicke 1997; Stern, 1985, 2004; Zahavi, 2000, 2004, 2005). While only some of these works deal directly with speech and sound, all, however, explore human experience in an attempt to retrieve it from the lethal reductionism of positivism, cognitivism, deconstruction and postmodernism, restoring, to some extent, the subject’s agency and individuality.

I.6 Chapter Outline

In the present chapter, I have positioned sound in the psychoanalytic clinical arena as the subject matter of this study and consequently described the unique dimensions it creates. The significature of nonlinguistic sound has been proposed as
the central research question to be elucidated. The disciplines for inclusion in this inquiry – psychoanalysis, developmental research and music theory – and their relation to sound have been addressed. Next, I propose a conceptual infrastructure for conducting the present investigation.

In chapter II – A conceptual context: Sound in Phenomenology, Music Semiotics, and Dialogism - I will explicate the three theories that coalesce into a conceptual matrix in the investigation of the nonlinguistic sonic dimension.

Chapter III – The Voice in Psychoanalysis - is devoted to the subject of the nonlinguistic voice as sonorous object and its meaning in the psychoanalytic clinical arena. The first part of the chapter consists of a critical survey of psychoanalytic literature on this subject. Various psychoanalytic conceptions of the voice will be addressed and critiqued. In the second part of the chapter, I present an alternative, multidimensional conception of the voice in psychoanalysis that integrates various levels of signification. In the third part of the chapter, I describe several vocal phenomena and categories that emerge in the nonlinguistic sonic dimension. These include three operational vocal principles that characterize the sonic dimension: the principle of correspondence, the principle of repetition and change and the principle of narrative. Subsequently, I define six vocal configurations that, as I suggest, can be discerned in the nonlinguistic vocal flow. These are: vocal timbre, vocal gesture, pause and silence, vocal persona, vocal duet and vocal ambience. The form and nature of these phenomena will be described and explained.

In Chapter IV – Listening - I address the meaning of listening to the nonlinguistic sonic dimension. Since listening is considered in this study to be the overarching principle underlying the total sonic dimension, and since it will have been addressed
in several manners in previous chapters, this last chapter focuses on the possibility of a *musical* listening in the psychoanalytic framework. I first consider psychoanalytic literature on listening through a concise survey, the main critique being that, though permeating the whole psychoanalytic operation, psychoanalytic listening is basically linguistic and therefore partial and reductive. Through the elucidation of the recent use of musical metaphors in psychoanalytic literature, I posit listening as a primarily embodied, sensuous prereflective experience emphasizing its essentially participatory nature. However, the relevance of music to illuminate the psychoanalytic clinical situation is not only metaphorical, and is related centrally to its ability to represent emotions in its own terms, as elucidated by the aesthetician Suzanne Langer (1970) and the music philosopher Victor Zuckerkandl (1957). In the second part of the chapter, listening is addressed through the summary and conclusion of the entire study. The course of the study will be evaluated in terms of findings and future directions for research. Inspired by the thinking of Zuckerkandl (ibid.), I conclude with the idea that a musical listening to the nonlinguistic sonic realm in psychoanalysis opens before us a different notion of subjectivity, intersubjectivity, time and space than the reified visual one we are accustomed to. Through this notion, music emerges, beyond its role as aesthetic object, as a fluid, participatory, interpenetrating way of knowing the lived-world, providing a different mode of approaching the nature of lived experience, and a powerful metaphor for its articulation.
Chapter II
A Conceptual Context: Sound in Phenomenology, Music Semiotics, and Dialogic Theory

II.1 Introduction

The present chapter proposes a conceptual framework created through the interaction of three theoretical perspectives on sound: Phenomenology, music semiotics and Mikhail Bakhtin's theory of dialogism. As explained in chapter I, the conceptual framework that emerges will facilitate the investigation of three central features of the sonic dimension: a. Sound's form, or, the continuous, dynamic, evolving nature of sonic phenomena; b. Sound's content, or, the ways sound creates meaning in light of its abstract, nondiscursive, processual structure, and c. Sound's presentation, or, the fundamentally experiential (embodied) and performative character of the sonic.

After elucidating the relevant concepts drawn from each theory, I will indicate their commensurability. The interrelationship, common themes and concerns between the three will be pointed out. These commonalities and meeting points are claimed to create a shared ground through which nonlinguistic sonic phenomena in speech can be more richly conceptualized. An example from the clinical psychoanalytic arena will illustrate how the suggested theoretical context works.

Finally, I indicate the relation between these theoretical principles and the contemporary scholarship on development, neuro- and cognitive science.
II.2 A Phenomenological Approach to Sound

II.1.1 Discovering the Nature of Sonic Phenomena

The first goal of a phenomenological approach to sonic experience is to move away from common assumptions and implicit beliefs about sound. It strives to create a “naïve” experience of the sonic and sets out to discern and carefully describe how sonic phenomena is given, in other words, how it is perceived and experienced by the individual listener. As a consequence of this phenomenological stance, it is expected that new facets of the phenomena of listening and voice will emerge.

A phenomenological move toward sound in speech results in the discovery of new and richer manifestations of this phenomenon than were previously thought to exist. A common effect of any phenomenological investigation is that it reveals that our language and thought about the phenomenon, sound in this case, is biased by traditions that conceal some of its facets. For example, it is a commonplace to treat sound as a purely abstract, temporal phenomenon. Yet, a close phenomenological examination reveals that sound has implied spatial aspects, such as materiality (as when a sound reveals the material or volume of the object that emitted the sound) directionality (as when we sense that a sound is coming toward us, or when we feel we are invaded by, or immersed in sound).

Of central interest to the present study is the fact that vocal sound is perceived through material, textural metaphors - as when a voice is felt to be “rough” or “velvety”. A phenomenological investigation of sonic experience reveals linguistic habits about sound that hide its true nature: for example, sounds are conceived as
high or low, and music is thought of as a moving object (in both examples, spatial attributions to sound are metaphoric. Some of the reasons why these spatial conceptions about sound developed will be discussed in chapter III).

These spatial and corporeal mappings onto sound need to be understood. Phenomenological examinations of these concepts have been conducted for example by the phenomenologist Don Ihde (1971, 2007) and the music philosopher Victor Zuckerkandl (1957), revealing the biases in existing linguistic traditions and, importantly, the way that the human body with its own corporeal schemata looms large in sonic perception. As a consequence of the suggested phenomenological outlook, it is anticipated that a new language about sound will emerge in the course of the present study.

In accordance to the phenomenological method, a distinction needs to be made between two elements that appear in ordinary experience as one: a. the sonic phenomena themselves and b. the process of listening or attending to sonic phenomena. This distinction has been drawn in relation to Husserl’s important concept of intentionality (this concept refers to the way consciousness is directed towards objects. It was developed by Husserl after Brentano and had a wide influence continental philosophy in the twentieth century). The concept of intentionality was later expanded in Heidegger’s concept of Dasein (1927/1996), and Merleau-Ponty’s idea of “being-in-the-world” (1962/2008). In brief, these authors maintain that the process of perception and its object are always experienced as one. Yet their systematic investigation implies a methodical separation between

---

23 Lakoff and Johnson (1980) have importantly argued the body to be central in the generation of all linguistic metaphors.
objects of perception and processes of perception or apperception of the phenomena. These two elements of sonic experience will be accordingly discussed in separate in chapter III and IV.

The main problem in the phenomenological approach, according to Ihde (2007, and see also 1971), is that none of us are any longer naïve listeners. We are always-already immersed in thought formed by verbal language. Furthermore, as I will demonstrate later on, a phenomenological elucidation of nonlinguistic sound within a linguistic happening, such as the psychoanalytic conversation, increases even more the difficulty of listening naïvely. It is only by a “second naïveté”, claims Ihde, that we can we approach a fresh listening. This second naïveté comes only by concentration and a willingness to suspend our preconceived notions of the phenomenon in question. This move represents phenomenology’s central method developed by Husserl, called “epoché” or, bracketing. An epoché consists of a first step in a phenomenological approach to a phenomenon and strives to suspend or step back from our ordinary ways of looking at things, setting aside our usual assumptions. Martin Heidegger, using Husserl’s philosophy as a methodology to formulate his ontology of being, held that the only way to get to the essential in things, is by “letting them be” or by “letting them show themselves” (in Ihde, 1971). In the same vein, and similarly basing herself on the thinking of Heidegger and Gadamer, the philosopher and psychoanalyst Gemma Corradi-Fiumara (1990), formulates an entire philosophy of listening, and similarly strives to recuperate the capacity of listening in a predominantly logocratic culture. This approach is evocative of Freud's original concept of analytic attention that he termed “evenly
suspended attention” (1912). In this seminal paper on analytic listening, Freud advocates a stance of listening that is reminiscent of the phenomenological bracketing of preconception. However, Freud was motivated by a powerful theoretical preconception - his theory of the unconscious. Therefore, his waiting for “what shows” was in fact preordained. Consequently, the psychoanalytic movement grew even further away from phenomenological intuition.

At the outset of this investigation, a major caveat is in order. As a rule, experience is total, multi-sensory and chaotic (Stern describes it in 1992 as an experiential “pandemonium” that the individual strives to organize and regulate), and impacts us through multiple modes of perception. In this context, listening is not an exclusively aural activity. We have already discussed how sonic experience contains implied spatial aspects, and how sound imparts upon hearing a sense of materiality, texture, volume and distance. In addition, our apprehension of sound is often accompanied by visual impressions. The conception of the senses as separate, distinct channels of experience stems from a philosophical-scientific tradition that has evolved in Western culture since the early Greeks, answering to culturally determined priorities of categorization and the need to create distinct, common, shareable ways of

24 It is interesting that Merleau-Ponty names Freud in relation to phenomenology. He asserts in his *Phenomenology of Perception*, that “phenomenology existed as a movement before arriving at complete awareness of itself as a philosophy..” and counts among its precursors Hegel ,Kierkegaard, Marx, Nietzsche and Freud. (1962, p. viii).

25 However, in the last decades, the influence of the linguistic turn and deconstruction brought psychoanalysis to a growing awareness of its philosophical implicit preconceptions. Moreover, a psychoanalytic interest in understanding this philosophical embeddedness seems of late to be on the rise (see for example Benjamin, 1988; Frie, 1999, 2001; Gendlin, 2003; Mitchell, 1993; Priet, 1999; Stolorow, 2003; Thompson, 2000). As I will discuss in chapter IV, this has greatly affected psychoanalytic listening.
dealing with experience. Inevitably, this categorization takes place with a simultaneous reduction and loss of the richness and nuance of global experience (Ihde, 2007). In his theories about young infant experience, Daniel Stern (1985) posits a similar idea about the reduction of experience in ontogeny brought about by the necessary creation of distinct, shareable categories, especially by language. In effect, a focusing on listening as a specific sense-experience (auditory) is an exercise carried out for analytical purposes, yet it is taken up as a tactical move to advance a deeper understanding of sound.

Because the present study takes place in the confines of a conversational situation, it is imperative, from a phenomenological standpoint, to temporarily suspend previously established presuppositions and categories about sound’s operation within speech. The first and strongest preconception in this case subjugates the sonic to semantics and regards the voice in the exclusive service of verbal discourse. Those vocal expressions that exceed language and do not directly serve linguistic utterance formation (vocal tone, tempo, stress and pause), are designated as "para-linguistic" – a term that reaffirms vocal sound’s subordination to verbal language, as I have argued in chapter I. Several conceptualizations of the nonlinguistic aspect of voice and speech have been made, the most infamous perhaps being Saussure’s (1916/1983) parole (spoken language) versus langue (language as abstract structure). By perpetuating the binaries word/sound, verbal/nonverbal, language/speech, speaking/writing, these conceptualizations consistently relegate nonlinguistic sound and the performance of speech to the weaker pole of the binary, to a subsidiary role, and, especially in the structuralist linguistic tradition were
considered to be a surplus (Dolar, 2006). These dichotomous definitions disclose, in Foucaultian terms (1970), the growing power and authority of word (and writing) that evolved throughout 17th century, culminating in the linguistic turn and deconstruction: Derrida’s conception of the written word (1997) based on Rousseau’s treatise on the speaking voice (1781/2008) effaced the impact of the voice even further.

Against these logocentric views, I pursue the idea that a powerful nonlinguistic sonic dimension is at work in speech as a partially autonomous semiosis that signifies specific content while interacting with the linguistic level. The operation of the nonlinguistic level of signification is different than that of the linguistic, in terms of both structure and principles of operation. This distinction and potential opposition between different systems of signification in the voice allow the understanding of multidimensional structures such as irony and double-messages, the poetics of false and authentic speech, and the nuanced qualification of the more arbitrary linguistic symbols. A phenomenological exploration of sound demonstrates the unique and autonomous action of nonlinguistic sound, existing simultaneously with its linguistic function in the creation of discursive meaning. Thus, Ihde advocates that in a phenomenology of sound

a distinction must be made that specifically distinguishes the ‘linguistic’ form of language from ‘language’ as the significant. ‘Linguistic’ language is language-as-word. It is the center, but not the entirety of language in the broad sense...The center of language as language-as-word may be understood as similar to the appearance of a focus-fringe phenomenon in the
sense that deployed around language-as-word is a vast field of meaningful activities which may in the broader sense of language be called ‘languages’. These ‘languages’ are gravitationally weighted toward the central significance in word, but they may be relatively distinguished from the ‘linguistic’ form of language... (Ihde, 2007, pp. 147-148)

A phenomenological approach reveals that sound “speaks many languages” simultaneously. In its nonlinguistic capacity (through tone, pause and rhythmic pattern), it purports to express meanings that language-as-word is much more limited of generating (like atmosphere, character, irony and myriad other qualitative nuances and emotional shades that language cannot name). These contents often do not figure at all in the verbal text. Furthermore, as has been demonstrated by infant research, it is through vocal tone, rhythm and gesture that an interpersonal, intersubjective sympathetic communion is created.

An illustration of the autonomous function of nonlinguistic sound in dialogue is the phenomenon of vocal ambience: One of sound's fundamental feats in dialogue is the constitution of an affective mood that is constituted sonically beyond (though simultaneous with) the verbal plane. The continually evolving vocal ambience is organized in sonic quasi-musical ways that are structurally different than the

26 Moreover, in writing, which in psychoanalysis is a central mode of reporting, transmission and discourse, the sonic nonverbal dimension all but disappears. This constitutes a central problem in psychoanalysis: the psychoanalytic encounter as a living event, occurring through voice and listening, is reduced through analytic practices of transcription that eliminate the sonic, living quality of the psychoanalytic experience from consciousness.
linguistic structure of the conversation. A quite banal example is a conversation that, in terms of its linguistic semantics, seems positive. However, the conversation unfolds in an admittedly gloomy atmosphere (signaled by the dialogue partners' tone of voice, facial and gestural parameters, rhythm and tempo). This double-featured discourse is not only about two things, but is performed, simultaneously, in two different manners. As opposed to the linear, directional speech structure, where utterances are relatively clear, stable (repeatable) and deployed in a reasonably ordered form of turn-taking, the creation of ambience occurs out of awareness, constituted simultaneously by both dialogic partners via various vocal mechanisms such as mutual correspondence and regulation. The unfolding of vocal ambience is continual, nonlinear, gradient, polyphonic, and multivalent. It is of interest to note that, in the case of ambience, the interlocutors' experience is twofold: not only do they create the vocal ambience, but they experience being immersed, contained and affected by it throughout the encounter. In this case nonlinguistic sound acts as both a means of articulation and as a container in the creation of a sonic surrounding “medium”.  

II.1.2. Oscillating Attention and Global Experience

Another relevance to using a phenomenological approach to the study of sound has to do with a certain resonance or affinity between the structure of general experience and the particularity of sonic experience. Being temporal and in constant flux, sonic phenomena (musical or not) are experienced as a continual stream

27 The phenomenon of vocal ambience will be addressed in more detail in the next chapter and will be claimed to be a higher-level phenomenon in a hierarchy of vocal events smaller, lower level sonic events, like vocal timbre and vocal gesture.
occurring in the here-and-now, its patterns and coherences unfolding in a continual flux, thus exemplifying the structure and unfolding of all experience. As has been proposed by the philosopher Suzanne Langer (1970), in this sense music operates as a sonic symbol of emotional, sentient experience and the ways it unfolds.

Phenomenology, as the study of experience and phenomena (things as they appear in our experience), with its particular concern for consciousness and the way we experience these phenomena and their meanings, is particularly apt to the investigation of sound. Not only does it provide conceptual and methodological tools to deal with specific features of sound, but it also facilitates the move from a purely descriptive to a conceptual level of discourse.\(^{28}\)

An example of this affinity between general experience and sonic experience is the phenomenological observation that phenomena are given to our attention through a center-periphery structure. Perception is organized in such a way as to be able to locate certain objects and focus on them, while simultaneously maintaining a more diffuse, not necessarily conscious attention to the field surrounding the object. The perceptual field (this is clearer in the visual field than in the auditory) has a certain limit, beyond which nothing else can be perceived. This limit is called in phenomenology a “horizon”.

Ihde gives the following example, indicating that a close attention to the sonic experience reveals that as a listener, one constantly experiences a multitude of sounds.

\(^{28}\) Curiously, phenomenological approaches to music are scarce (exceptions are Clifton, 1983, Zuckerkandl, 1957).
Thus, I hear at one time the hum of my computer, the clicking of the keyboard as my fingers type, the sounds in the other room made by my wife, cars and birds outside, the humming of my own nervous system, the sounds of music coming from the radio and so on. I can direct my attention at will to any of these sounds, and automatically place all the other sounds in the periphery of my attention. In fact, the phenomenon of active attention is about a constant directing and redirecting my focus of attention to different phenomena in the field... I can sustain my attention on a certain phenomenon for a period of time, but at some point my attention will be directed elsewhere, causing a shift in the relationship of center-periphery within the sonic field. (Ihde, 2007, p. 51).

The ability to consciously and voluntarily shift our attention (though attention shifts also involuntarily, something that will be addressed later) indicates that there is no one fixed, objective, ontological center of sonic occurrence, but that the center is always relative and shifting. Similarly, while listening to music, one can focus on the sound of the violins in a symphony and be able to follow their melody, without losing the perception of the sonic field surrounding the violins (though this perception becomes, as it were, more blurred). Then I can shift my attention to the sounds of the flute and the violins will then “move out” to the periphery. Attention can thus be thought of as a series of fluctuations or oscillations between changing

---

29 The concept of a relative, shifting center that determines a periphery (context) is a key principle in the philosophy of Bakhtin that will be addressed later in this chapter.
core-periphery structures. Shifting attention delineates, by its very motion, a multidimensional field of experience that is nonlinear and multidimensional.\textsuperscript{30}

And yet, the nature of auditory perception is such that we are also able to take in the field as a whole, and listen to the entire gestalt of present sounds as an entity in itself, similar to the way one visually perceives a landscape. This soundscape is not static, rather it is a dynamic entity that changes along with its constituent sonic occurrences.

These phenomena, that are true of all perception, are quintessential in the field of sound and characterize the dual nature of listening. On one hand, we have the ability to focus on specific events in the sonic field and direct our attention at any given moment to varying phenomena, such as the sounds of the violins or the flute. In conversation we similarly can focus on the meaning of the interlocutor's words, the tone of her voice or her specific manner of stressing the ends of sentences. On the other, we have the ability to perceive sonic phenomena globally: we can listen to the symphony as a whole, and while participating in a conversation we can experience its global sonic features through phenomena like ambience. Thus there is a series of attention fluctuations between as-a-whole (global) and local levels of events, operating to create an interpretive structure, in other words, to generate meanings. For example, attention processes are quick to establish phenomenal

\textsuperscript{30} A similar process in relation to vision has been described by the psychoanalyst Marion Milner (1987).
invariants, that in turn set up sonic norms (expectations or stable patterns) through which new occurrences can be discerned and assessed.  

Thus, listening consists of an ongoing dialogue, or shifting, between a sweeping, total experience of sound (the gestalt of sonic experiencing) and a more focused, pinpointed attention to some factor in the sonic field. In other words, it can be stated that there is interplay between unity and plurality in sonic experience. For I perceive my interlocutor’s utterances as both a unity of utterance (unity of literal meanings with quality of utterance and performance) and yet I can pluralize the utterance and discern in it several layers of expression, focusing on a particular feature and making it stand out in my attention. Sonic experience exemplifies the fact that there is always an interaction between experience on a higher, more generalized level, and a lower, local one (or macro- and micro-levels). Hence, Freud’s original concept of “evenly suspended attention” (1912) can be understood as a form of macro listening - a sweeping stance of attention that places no focus on anything in particular and yet interacts with a type of local, more precise, focused attention on certain words (as Freud proposed through the metaphor of the analyst-listener as surgeon in the same text). The oscillation between different modes of listening was consequently elucidated quite extensively in the analytic literature (see Brenneis, 1994; Carlson, 2002; Freedman, et al. 1974; Freedman, 1983; Gardner, 1991; Ogden, 1999a; Reik, 1958). But although these writers have observed and described the oscillatory nature of listening, it was largely directed at the linguistic level of happening. In addition, only few writers have understood its

---

31These processes have been elucidated in music theory by authors like Meyer (1956), and in infant research by Stern (1985).
deeper implications (one such writer is the psychoanalyst Marion Milner, [1987]). In the field of music however, these implications were much discussed, for example by Zuckerkandl, (1957) and HaCohen(2001a, 2001b). Auditory oscillation is particularly revelatory of the diffuseness of boundaries and horizons and the interpenetration and simultaneity of phenomena, as opposed to the relative solidity and distinctness of the visual sphere.

The actual structure of sonic perception resonates with my present view that every sonic phenomenon (and its attendant listening process) is situated, in other words, part of a complex, multi-dimensional, continually changing context. Accordingly, human interaction and communication can be apprehended as a complex event wherein several semiotic systems of signification - linguistic, paralinguistic, nonlinguistic (as well as gestural and facial, though these are not discussed here) – interact simultaneously, rendering human communication a polyphonic and intertextual event. Verbal language, with its symbolic, more arbitrary and unmotivated structure, generates meanings that are more univalent (hence, emphasizing the dimension of unity). At the same time however, it interacts with the nonlinguistic vocal occurrences whose meanings are characteristically motivated, indeterminate and multivalent. Attending to these several foci of meaning results in the creation of a nonlinear, multidimensional space where meanings are generated through the dialogization of different forms of discourse or different semioses. This stance calls for the cultivation of a shifting, oscillating, contextualizing attention – precisely that which is privileged by the phenomenological operation. It is not
coincidental that oscillating attention is characteristic of musicians and its promotion is high on the agenda of musical ear training. Musical forms of listening are certainly related to sound’s nondiscursive nature since nonlinguistic sound is largely nonpropositional and does not have direct external reference, hence, one of the central ways of understanding its meaning is through the internal relations between the different components of the sonic system.  

If we concentrate on the aspect of experiencing sound through an oscillation between global and focused listening, and through a motion between several core-periphery structures, we can claim that the listener, through her organization of the sonic phenomena, her conscious and unconscious choices of core-periphery structures and her dialogization of such structures and elements, actively “orchestrates” her auditory field. This idea implies that selective attention and modes of applying attention to speech and the sonic nonlinguistic dimension, constitute a creative activity that employs auditory stimuli to generate meanings. Part of this creative act is undoubtedly determined by social convention but another part is created by the listener’s individuality and idiosyncrasies.

This of course touches upon the vast problem of interpretation. In this regard, the idea of oscillating attention suggests that, in the psychoanalytic situation, the continual shifting between foci of attention and phenomena that become thematized, creates a trajectory. By directing her attention (voluntarily and

---

32 I do not mean to imply that meanings are generated only through internal relations within the sonic system, since it is also through interaction with the surrounding culture that the hermeneutics of sound is constituted, as has been importantly demonstrated by recent studies in music and culture (the literature on this subject is vast, but see for example Kramer, 2002; Shepherd and Wicke, 1997). Nevertheless, inner relations and formal analysis continue to be important measures in the understanding of sonic fields.
involuntarily) to the interlocutor’s verbal discourse, to her tone of voice, her rhythm, body language, her own silent stream of thoughts and associations, her own sounds and her bodily sensations, the listener’s attention actively creates, or as I have suggested, orchestrates the auditory space. Differently than the visual space, which is more clearly given to the observer, the auditory space is creatively (and thus subjectively) generated by the listening activity.\textsuperscript{33} From this point of view, it is important to understand how one goes about this oscillation (this will be elaborated in later in this chapter, through Bakhtin’s concept of dialogization). For example, what are the listener’s varying foci of attention? What is selected by the listener as worth of attention? What is rejected or ignored? Which vocal sounds typically draw the listener’s unconscious attention by their specific timbre? In this light, more rigid or flexible styles of analytic listening could be understood as outcomes of personal styles of oscillation, representing further possibilities for analysts (and maybe also analysands) for self reflection.

Beyond global and focused listening, the structure of sound and music presents the field of perception with other manifestations of multi-dimensionality that, again, epitomize experience at large. Thus, music presents a temporal dimension of succession (sound after sound)\textsuperscript{34} and a synchronic dimension of simultaneity: the ability to hear two or more sounds at once. Sometimes this happens in a concrete manner, as when partners enter each other’s speech. Other common forms of

\textsuperscript{33} Of course this is true also of visual activity, however, in the auditory realm this interpretive gesture is much more pronounced and fundamental.

\textsuperscript{34} A succession that is nevertheless apprehended as a gestalt form. When we hear sounds coming from a flute, for example, we hear them as melody, not as a succession of discrete sounds. This holistic nature of perception was addressed not only by phenomenologists, but also for example by the philosopher Henri Bergson (2001/1913), as well as by the developmentalist Daniel Stern (1985, 1992).
simultaneity in psychoanalysis is the multivocality of outer and inner speech and, as I have indicated above, the polyphony of several semiotic circles operating all at once.\textsuperscript{35} Another polyphonic manifestation, that will be better elucidated in the next chapter, is the sonic coalescing of past and present. I suggest that past occurrences can be discerned in the interlocutor’s vocal utterance in the present. For example, a characteristic timid manner of speech may indicate a historic relational pattern that was internalized and is manifested in present speech. Thus, the past, ordinarily thought of as something bygone, materializes in present occurrences.\textsuperscript{36} In each of these dimensions (succession and simultaneity), there is a possibility to listen in a focused or global way. These continua of center-periphery, unity-plurality, succession-simultaneity, linearity-nonlinearity, flux-stability, are quintessential of sonic phenomena and its corresponding mechanisms of perception, and will be more closely examined as they appear in the psychoanalytic clinical encounter in chapter III.

\textbf{II. 1. 3. Embodied Consciousness}

A third reason for using a phenomenological approach and method is its interest in present time experience and its elucidation as embodied. This outlook has a particular relevance to sonic experience and to vocal-auditory communication. As voice and listening inhere primarily in the body, their operation through expressive

\textsuperscript{35} The idea of the analytic encounter as polyphony or symphony, i.e., as a texture composed between interacting voices has been proposed by several psychoanalysts. See for example, Benjamin, (2002); Bollas, (2007); Knoblauch, (2000); Mitchell, (1993); Priel (1999, 2003).

\textsuperscript{36} Of course, this linear conception of time is another characteristic of positivistic and visuo-linguistic attitudes. However, the condensation of past (and future) into the present is a phenomenon evident in music, and discussed as such by Bergson, (2001/1913), by the phenomenologist tradition and, in their footsteps, by Stern (2004).
tone, vocal gestures and rhythms indicate the operation of an embodied and sentient emotional intelligence. The ideas of the phenomenological philosopher Maurice Merleau-Ponty are of special relevance here. According to Merleau-Ponty (1962), the center of experience is, indeed, the body. Merleau-Ponty's thinking argues against the overlooking of the centrality of the body in human experience by the philosophical tradition. Merleau-Ponty claims that bodies live in the world and understand it through their immediate, embodied and emotionally intelligent connections with it. The experiencing subject is not located outside of the experienced world, nor screens it through cognitive, symbolic or objectifying functions. Rather, the subject is as Merleau-Ponty phrased it, “in-the-world”, that is, located in the world and experiencing it directly through the body as part of the world. The subject's incarnate experiences are, in Merleau-Ponty's terms, “lived-experiences”.

Merleau-Ponty proposes the concept of the “body-subject” as an intelligent, holistic entity that directs its behaviors in a fluid, integrative fashion, creating an attuned and corresponding relation between self and environment, both of which are conceived as being of the same fabric. It is the holistic body-subject in its total experiential capacity that understands its world - not an abstract mind that inhabits the body. The concept of the body-subject conceives the relations between body and world as a fundamentally incarnated experience, without privileging either mental or material aspects of that experience, thus rejecting the Cartesian mind-body dualism. Merleau-Ponty elucidates the ambiguous nature of the body and views bodily existence as a third category: “The lived body is neither spirit nor
nature, neither soul nor body, neither inner nor outer, neither subject nor object.” (in Gallagher and Zahavi, 2008, p. 136).

Regarding the interpersonal realm, Merleau-Ponty's radical explanation of the mutual sensuous recognition of individuals similarly emphasizes that it is the body that encounters others in the world, not an abstracted mind that inhabits that body. The body-subject's lived experience is necessarily located and its language is one of sounds, gestures, movements and action. Thus, the encounter with others in the world is primarily an embodied event that conditions the very possibility of communicating. In *The Phenomenology of Perception*, Merleau-Ponty states:

> ... it is precisely my body which perceives the body of another person, and discovers in that other body a miraculous prolongation of my own intentions, a familiar way of dealing with the world. Henceforth, as the parts of my body together comprise a system, so my body and the other person's are one whole, two sides of one and the same phenomenon.....in the experience of dialogue, there is constituted between the other person and myself a common ground; my thought and his are interwoven into a single fabric, my words and those of my interlocutor are called forth by the state of the discussion, and they are inserted into a shared operation of which neither of us is the creator. We have here a dual being, where the other is for me no longer a mere bit of behavior in my transcendental field, nor I in his; we are collaborators for each other in consummate reciprocity. Our perspectives merge into each other, and we co-exist through a common world. (1962/2008, p. 354).
In the present account, voice and listening are central components of interpersonal, reciprocal lived experience; an embodied meeting between humans that, through an innate sensing of each other's sounds, rhythms and expressive gestures, constitute a sympathetic, intelligent intersubjective encounter. Merleau-Ponty's conception of language liberates it from rigorous structuralist conceptions (such as the Saussurian model) and the nullifying effects of deconstruction on the subject's voice and agency, and reclaims a performative, lived, agentic speech:

To understand speech, we do not have to consult some inner lexicon which gives us the pure thoughts covered up by the words of forms we are perceiving; we only have to lend ourselves to its life, to its movement of differentiation and articulations, and to its eloquent gestures. (Merleau-Ponty, 1964, p. 42).

The communicative activity in which human beings engage comes about through the expressiveness of their bodies: by facial expressions, gestures, body-language and styles of behavior, vocal tones and gestures, variations of pause and rhythm. According to Matthews (1992), Merleau-Ponty's considers human expression to be a transformation rather than a representation. Thus, in language we do not simply reproduce what is in the world –outer and inner alike - but transform it into an expression. This does not mean that our minds impose a meaning on something, but rather that meaning emerges from the contact between the subject and a world of which, the subject is a part. For Merleau-Ponty, the seemingly arbitrary form of the word would no longer appear so, if we would take into account the emotional content of the word, its gestural sense.
...it would then be found that the words, vowels and phonemes are so many ways of ‘singing’ the world, and that their function is to represent things not, as the naïve onomatopoetic theory had it, by reason of an objective resemblance, but because they extract, and literally express, their emotional essence…..(representing) ways for the human body to sing the world’s praises and in the last resort to live it. (1962, p. 187).

These communicative capabilities, considered by traditional views to be pre-verbal (in a sense that they are ontogenetically earlier and thus more primitive; discarded after verbal skills are acquired) are here viewed, in the spirit of Merleau-Ponty, rather as a fundamental matrix of communication out of which word emerges. Yet this matrix always-already surrounds word, permeates it and gives it life, and, as I will argue, is no less than the essential means for the constitution of a shared, consummate reciprocity. From this point of view, the dichotomy of verbal-preverbal needs to be reunited within speech and communication conceived in a more holistic manner. Although language-as-word is largely the center of verbal communication, the sonic nonlinguistic dimension permeates and envelops it like an aura that imparts life and uniqueness to each and every utterance.

The above outlook postulates the analytic conversation as primarily an embodied, situated encounter, whose meanings emerge from the corporeal, sentient experiences of self and other. In this encounter, an intersubjective, sympathetic experience is constituted whereby the other is sensed and known initially not

37 This conception is strikingly similar to Rousseau’s theory on the origin of languages (1781/2008).
through the literal meaning of discourse, but through a corporeality that impresses itself via the body’s rhythms and gestures, voice and mutual listening. This has, in my opinion, a radical meaning: beyond the establishment of a benign, sympathetic atmosphere, the embodied vocal interaction implies the co-creation of a dialogic semiosis, only a part of which is verbal (through different terms, the psychoanalyst and linguist Julia Kristeva also viewed the polyphony of the dyad's discourse [1986a, 1986b) and its semiotic order. The meeting of the other begins as a pre-reflective knowledge and ethic, not as a reflective interpretation of verbal discourse. Based on Merleau-Ponty's conception this encounter is here viewed as embedded in the subject’s own incarnated self-experience.

The conception of a body-subject and pre-reflective lived-experience is inevitably connected to the question of consciousness that is central, though in radically different ways to both phenomenology and psychoanalysis. According to the two phenomenological scholars of subjectivity, Gallagher and Zahavi (2008), the phenomenological view maintains that a minimal form of self-consciousness is a constant structural feature of conscious experience. Experience happens for the experiencing subject in an immediate way and as part of this immediacy, it is implicitly marked as “my” experience. The immediate and first-person givenness nature of experiential phenomena must be accounted for in terms of a pre-reflective self-consciousness. In the most basic sense of the term, self-consciousness is not something that comes about the moment one attentively inspects or reflectively introspects one's experiences. Nor does it occur in the instant of self-recognition of one's image in the mirror, in the use of the first-person pronoun or in the
construction of a self-narrative. Rather, different kinds of self-consciousness predominant in classic psychoanalytic conceptions are to be distinguished from the pre-reflective self-consciousness that is present whenever one is living through or undergoing an experience, consciously perceiving the world, thinking an occurrent thought, feeling thirsty or happy, singing or listening to a melody. In the present study, voice and listening are considered as primarily occurring in this mode of pre-reflective consciousness. From this point of view, all reflective consciousness, including the highly cultivated, reflective psychoanalytic listening, is based upon this primary “listening in the flesh” and draws its meanings from this level.

Whether and how the conception of pre-reflective consciousness can be reconciled with the classical psychoanalytic unconscious is a question that the scope of the present study does not permit to address in full. Consciousness in psychoanalysis traditionally refers to the ego structure, while the classical repressed unconscious refers to the existence of mental ideas and thoughts that were once conscious and underwent later repression, thus existing, as it were, in a concealed manner in the subject's mind. Because psychoanalysis is built upon linguistic precepts, it refers to consciousness in its relation to language, thus, symbolic thought. In other words, consciousness in classical psychoanalytic theory is linguistic, reflective and occurs after the fact. It is a consciousness that can thematized, objectify, make reflections, produce narratives and make interpretations. It is only in the recent decades that

38 The concept of the repressed unconscious has been itself a major area of debate in psychoanalysis, challenged and revisited numerous times over the 20th century. This vast theme cannot be addressed here, but see for example Kennedy (1996).
psychoanalysis began to address other conceptions of consciousness and embodied experience.

II.3 Sonic Meaning and Signification: Sound in Music Semiotics

II.3.1 Introduction and Aims

The phenomenological stance described above facilitated an approximation to sound from within the vantage point of the experiencer. As we have seen, this first-person stance, both as a philosophical position and a methodical approach leads to a renewed encounter with the sonic order, allowing for new aspects of sound in speech to emerge, aspects that our habitual listening and speaking practices usually conceal.

Given that we have phenomenologically accepted the existence of a sonic nonlinguistic realm in speech, the investigation will now proceed towards the subject of sonic signification. For this to happen, we must first elucidate in what way sonic phenomena within the analytic conversational sphere generates meaning by elaborating ways to segment and articulate the continuous sonic stream. In turn, this articulation will help in discerning more complex sonic mechanisms, structures and their activity.

For this purpose I turn to music theory where the issue of sonic nonlinguistic meaning has been famously debated. Music theory’s discourse on meaning is suggested as an alternative explanation to sound in speech. From this discourse – especially music semiotics, but also aesthetics- certain ideas and conceptualizations
are drawn that seem particularly apt for mapping on to the analytical conversational situation. More specifically, I employ discourse that conceptualizes musical gesture and posture, musical personae, sonic mood and ambience, voice and sonic subjectivity, musical narrative and musical metaphor. These concepts have been elaborated in the past three decades by music scholars including: Abbate, 1991; Cone, 1974; Cumming, 1996, 1997; 2001; Echard, 1999, 2000, 2005; HaCohen, 1999, 2001a, 2001b; Hatten, 1994; Lidov, 1987, 1999, 2005; Meyer, 1956; Monelle, 1992; and Tarasti, 1994. In the next chapters, these categories will be delved into in greater detail, while here, I will deal in a more abstract manner with their theoretical justification.

My aim is first to demonstrate that the sonic nonlinguistic realm can be - in a concrete, material manner- told apart from sound-as-word, and that articulations of this realm can be made in the process of listening. My methodological purpose is to advance phenomenological articulations and categories that are useable in vivo, in the course of the analytic encounter. This pragmatic method stands in contrast to empirical approaches and analyses of sound and the voice that take place after the fact, like text analysis (see for example Pittenger, Hockett and Danehy, 1960), or analyses of spectrograms and recordings. These formal analyses are traditional in music theory, and much in use in the clinical practice of music therapy. Though these analyses are clearly contributive to the understanding of sound's operation, my claim here is that listening to sound is an interpretive experience naturally made in the course of experience by conversational partners, in an unaware and
prereflective manner. It is these processes that are of concern in the present study, and shedding light on them will presumably lead to a heightened awareness of these in real clinical time. As we shall see, categories are designated that are predominantly body-oriented and metaphoric (the rationale of which will be explained later).

Secondly, I want to elaborate that, as in language, these articulations operate in a coordinated, systematic manner, and that signification emerges from several levels of interpretation— from local levels through basic, molar sonic units (like vocal timbre or gesture), to higher levels of varying complexity (like vocal mood and vocal persona), where basic categories combine and synthesize. However, I would like to note that I do not intend to posit the sonic dimension as a thoroughly closed system. Neither do I offer a tight methodological procedure for its elucidation. Rather a general interpretive orientation is suggested. It will be remembered that the sonic nonlinguistic dimension is only partially autonomous, as it interacts continuously with other ongoing levels of signification, like the verbal and the gestural. Sonic categories are just one pole (the material pole) of sonic semiosis. The other pole, as will be further explained, is dynamic and relational. Meaning is not found in sonic categories, but rather in processes of signification and interaction between sonic categories, verbal and gestural levels of signification. Analytic meaning is thus obtained through a polyphonic approach to a multivocal texture that is comprised of material elements (verbal, sonic, gestural) and their dynamic interaction, unfolding through concepts such as addressivity, energetic directionality, affective value, dialogic force and rhetoric intent (these concepts will be explained later on).
Thus, my third aim is to elucidate the dynamic pole of sonic semiosis. Here I suggest that meaning results from a multidimensional process whereby phenomena are experienced and understood as dynamic, polyphonic textures that continually unfold, intersecting affective and propositional, individual and social, material and processual, synchronic and diachronic dimensions. In their discussion about communication and meaning in music, ethnomusicologists Charles Keil and Steven Feld (1994) state:

Being fundamentally relational, communication is process, and our concern with it should be a concern with the operation of social determination in process. *The focus is always on a relationship, not on a thing or entity...* Communication in this sense is no longer ontologically reified as a transmission or force. It can only exist relationally, in between, at unions and intersections... Communication then is not located in the content communicated or the information transferred. At the same time it is not just the form of the content nor the stream of its conveyance. It is interactive, residing in dialectic relations between form and content, stream and information, code and message, culture and behavior production and reception, construction and interpretation. *Communication is neither the idea nor the action but the process of intersection whereby objects and events are, through the work of social actors, rendered meaningful or not.* (Keil and Feld, 1994, p. 78, italics are mine).

While Keil and Feld stress social determination (and psychoanalytic authors are mostly concerned about constructs such as the self and the unconscious) their basic
approach to understanding sonic meaning as a “process of intersection between objects and events” is essential.

**II.3.2 Sound, Music and Signification**

Since the scope of this study does not permit me to enter the subject of musical meaning in all its vastness, I will briefly address aspects of musical meaning that are most relevant to this research.³⁹

Music’s nonreferential essence gives rise to many questions regarding its meaning. As concisely expounded by Echard (1999), the problem of meaning in music has been at the center of debate for at least two centuries, leading to the formation of several declared positions the common denominator of which is that they all use linguistics as a model. An extreme non-referentialist position claims that music does not signify in linguistic, discursive ways (for example Hanslick 1954/1986) and that any affective or referential perceptions of it should be disregarded; a position that, as Echard is quick to point out, does not have a particular interest to musical semiotics (nor to psychoanalysis, I must add,). Other positions include the idea that musical events are significant only in their internal (formal) relations to each other (for example Meyer, 1957). A more referential claim posits that certain musical configurations can acquire fairly stable, consistent and shared connotations within a particular context, in effect becoming nearly as referential as lexical items in language (Ratner, 1980). In response to these linguistic based theories, other, phenomenologically oriented theories were formulated by authors such as Zuckerkandl (1957) and

³⁹ The reading on musical meaning is massive. Some important texts are suggested: Hatten, 1994; Kivy, 1980; Kramer, 2002; Langer, 1970; Meyer, 1950; Middleton, 1990; Shepherd and Wicke, 1997; Zuckerkandl, 1957.
Langer (1953, 1970), that centrally positioned embodied experience and perception in the process of musical signification.

But in the last decades, other readings of musical works and practices have evolved, integrating views from music aesthetics, popular music studies and ethnomusicology. These readings are more hermeneutically oriented and integrate formal analysis with extra-musical considerations such as cultural context, performance and reception practices. Within this development, the interpenetration between music and culture (thus, verbal language) has been increasingly acknowledged and scrutinized. The commonly held construction, whereby music is extra-cultural and ineffable has been subjected in recent years to revision from musicologists as well as from linguists and aestheticians. While not completely dismissing the notion of music as unmediated and offering privileged access to the rhythms of our body, music has been examined in its textuality and entanglement in language that must, of necessity, be used in processes of reception, response and dissemination. As a consequence of these relatively recent developments, musical meaning is largely conceived today as a nexus where aesthetic, cultural, historic, individual and social levels of signification are reflected in sound (see for example HaCohen, 2001b; Keil and Feld, 1994; Kramer, 2002; Shepherd and Wicke, 1997)

In a similar manner, music semiotics, a branch of musical studies that developed out of general semiotics, has made considerable progress in integrating hermeneutic and social levels of sonic signification with previous exclusively formal and structural
analyses. In addition to musical texts, recent music semiotics addresses musical events and musical actors, performance and reception processes as sign systems. (see for example Auslander, 2006; Echard, 2005; Monson, 1996).

One of the recent developments of music semiotics is the endeavor to develop nonlinguistic models of analysis (see especially Cumming, 2000; Lidov, 1999, 2005). Whereas older analytic methods relied almost entirely on linguistic models, some writers contend that, while structuralist views and methods can be well applied to music in explaining synchronic structures, forms and relations, they ultimately prove to be insufficient in elucidating musical signification. As a semiotic system, structuralism privileges binary, differential, non-motivated (convention-ruled) articulations, suggesting that sound as such is entirely supplanted by structure, having no significance in itself. However, while this claim may be somehow defensible in speech, in music it does not account for other types of sonic signs and effects that are motivated, concrete, and of a continuous, diffuse nature that nonetheless carry fundamental significance in musical experience. I similarly claim that structuralist linguistic analyses cannot account for expressive, emotionally nuanced signification generated through speech. The gap between language as structure and performed speech has indeed been recognized by Saussure, who, as we have seen, differentiates between langue, the abstract structure of language, and parole, speech, which he sees as merely a particular instantiation of language (and which he never set out to investigate). Thus, in a book dealing with the

---

interpenetration of music and language, the Canadian music semiotician David Lidov states:

Modern linguistics, as a science, seems generally to ignore the aspects of language which cannot be studied as systematic articulations. To say this is almost to say that linguists ignore what we might call the musical side of speech – the feeling tones and gestural character of speech. In part, the bias is the bias of writing, the prejudice of regarding what can be written as more essential to the medium than what cannot be. The scope of unwritable effects of speech, in maintaining authority, in asserting need, in communicating personal emotion, is in truth enormous. If these effects seem 'not of essence' in language, it may be also because compelling methods to study them are not in view. But there is another, better reason for the bias, the recognition of what language and only language can do. That special function has to do with particular types of reference...Those referential functions which can be fulfilled by writing and which are thus fulfilled by systematic articulations have a great claim to priority. Nevertheless, to regard this special capacity as the essence of language, as Saussure did, is logically indefensible, hides the density, prevalence, and human importance of other, expressively referential functions of language, and also needlessly handicaps the comparison of language and music. (Lidov, 2005, p. 3).  

---

41 Lidov’s statement, given from a music-semiotic angle is similar to Ihde’s phenomenological conception of language-as-word and its surrounding field. Bakhtin criticized the split between word and voice even more vehemently, as explained in the next section.
The music semiotic developments in the last decades have replaced the Saussurean model to a great extent, with the semiotic model of the American philosopher Charles Sanders Peirce for a theoretical base. Described below, Peirce provides a model that accounted for the whole of human activity as semiotic, including its manifold nonlinguistic manifestations. Peirce conceived of semiotic activity as a dynamic process, emphasizing the dynamic unfolding of signification (as opposed to the more static, synchronic model of Saussure). These two aspects of the Peircean model – the dynamic aspect and the ability to account for nonlinguistic signs - are naturally attractive to music theorists. The deeper consequence of the Peircean musical turn is that it represents a shift away from structuralist abstraction toward a more phenomenological interpretation of lived experience. Musical signs have begun to be interpreted as icons and indices of embodied experience, carrying their own expressive historicity, thus marking a difference from the Saussurean unmotivated conventional signs. The new conceptualizations highlight the uniqueness of musical signs as temporal, ever changing, processual signs, i.e., signs that develop on a temporal axis and unfold through time. These signs are not only different from the stable, synchronous, a-temporal linguistic signs, but imply different (also temporal) modes of apperception and interpretation.

In maintaining that musical experience unfolds temporally primarily through embodied schemas and metaphors, music theory increasingly posits the body and embodied experience as a central axis in musical experience and signification. These developments have opened the way to account for musical phenomena and

---

42 We have seen above that both phenomenology and certain psychoanalytic strands make this precise claim.
experience that characteristically resist verbalization, phenomena that led in the past to a mystification of musical experience (most famously, the creation of the musical trope of ineffability, see for example, Jankelevitch, 1983). Music is thus increasingly conceptualized through gradient, qualitative, energetic and affective axes where the body and related issues of affect and subjectivity are central, bringing about a long-awaited rapprochement between music and psychoanalytic theory (important psychoanalytic authors of this movement are Feder et al. 1990, 1993; Knoblauch, 2000; Lecourt, 1990; Noy, 1966; 1999; Rose, 1993, 2004; Schwartz, 1997; Stein, 1999, 2007). Notwithstanding the difference between the two fields, these authors pave the way for the relevance and feasibility of using musical concepts in psychoanalysis and address the vocal nonlinguistic dimension as a quasi-musical dimension that has both metaphorical and concrete values.

As fundamentally claimed in this research, my intention is to engage the inherent musical nature of the nonlinguistic dimension of the analytic conversation, in other words, to conceptualize it as music. Because of their continual, qualitative, flexible nature, the signifying activity of musical elements in speech is related to affective, energetic and relational levels of meaning. Like musical signs, sonic signs in speech are essentially presentational (as opposed to referential or representational), i.e. their meaning is not deferred from the sign, but is in the sign, emerging in the course of its performance, through sonic enactment. This makes them into a special kind of

---

43 In other words, I claim that in parallel to generating discursive meaning, sound generates affective and relational meanings through mechanisms other than linguistic. This "dual-coding" of language has been recognized and referred to by a number of psychoanalytic authors, as well as by scholars from the field of development and cognitive science, a partial list of whom includes Beebe, Knoblauch, Rustin & Sorter, 2003b; 1988; Bucci, 1985, 2001; Canestri, 2000; Knoblauch, 2000; Pally, 1996, 2001; Rizzuto, 2002; Rose, 2004; Stern, 1985; and Stein, 2007). Some of them have made explicit connections with music (especially Knoblauch, Rose and Stein (ibid.), while in others this connection is hinted at or implied.
sign, the nature of which was elaborated by Peirce. Although his model did not deal directly with music, his concepts, as mentioned above, have been attractive to music theorists. As evidenced by a large number of recent music semiotic conceptualizations, Peirce’s analyses of the interpreting subject and acts of interpretation and representation as continuously developing (rather than regarding signs and objects as static, “mindless” facts) has enormously enriched musical discourse. Since a large share of the musical concepts I use rely on Peirce’s model, and since his conception of the sign is essential for my understanding of sonic nonlinguistic events in the analytic scenario, the basics of Peirce’s ideas of the sign and their pertinence to sound and music will be described below.44

II.3.3 Peirce’s Structure of the Sign

For Peirce, as for Saussure, signs are relational rather than material entities. However, differing from Saussure, Peirce’s concept of the sign is triadic and consists of an obligatory relationship between three elements: an object, a sign, and an interpretant (which is his most innovative and important element and without which, according to Peirce, there would be no signification). While the sign (or representamen) parallels Saussure’s signifier45 and the object is similar to the signified, the interpretant is an element that acts to bridge or connect between the representamen and the object. A sign and its signified do not create signification in themselves, they need something to connect them and enable the work of signification. This something can be a process, a feeling, or a concept that takes

---

44 For a more extensive discussion of Peircian concepts and their application to music see Cumming, 2001; Echard, 1999; Hatten, (1994) and Lidov (1999, 2005).

45 For Saussure, as for most semiotic theorists, there is an essential dichotomy between signifier and signified. Saussure’s sign is a combination of the two, i.e., word and meaning. Both are not things, but ideas of things and positions within the linguistic system.
place in a person or in culture. This mediating object (in itself a sign) does not come out of nowhere but already exists in some way, as an element, an idea, a feeling, a system or process of relating sign to object. Relating smoke to fire, or an image of a man to a man, or the word dog to the animal, occurs through culturally existing interpretants. Because a sign can have many meanings, it can have many interpretants. The meaning lies not in the sign itself, but in the semiotic activity related to it. To take the above example, there is no consistent, stable meaning to “dog” as sign, since it can mean many things: it can be an example of a word, the idea of the animal, a word signifying fidelity, friendship, domesticity, or a class of domestic animals. Dog can signify a person (one who is at social disadvantage, as in “underdog”), a pet, a specific dog, or a situated dog - a dog signifies differently when gendered. A dog in Antarctica is not a dog in New York or in China or in Iraq; a dog in 2009 is not a dog in the middle ages or in Biblical times. The concepts and processes needed to determine which kind of dog one means are interpretants.

In terms of music, when I hear a piece of music as soft and comforting, I must have used “softness” and “comfort” as interpretants, i.e., as concepts of softness and comfort already existing in my mind – whether through embodied knowledge (my previous experiences of softness and comfort) or as conceptual knowledge (as ideas of softness and comfort that are qualified culturally and linguistically in a certain way). This previous knowing allows me to map this inner schema or interpretant onto the music I am hearing. It is important to note that the interpretant is mostly prereflective, although for purposes of analysis, it can become a conscious object of reflection. Peirce’s conception of the sign demonstrates that signs are biased and
that their signification is not fixed, but dependent on their being read as something (Echard, 2003).

The above conception of signification is, as we shall see, particularly germane to the fundamental features of sonic phenomena such as timbre, inflection, and intensity. These features resist verbalization because they are qualitative, gradient and fleeting – precisely those qualities that characterize the music of speech. Thus, in psychoanalysis, sonic phenomena such as vocal tone, timbre, gestural motif, pause, rhythmic pattern as well as higher and more complex sonic structures such as vocal persona and vocal ambience are attended to (though sometimes pre-reflectively, as discussed by Bucci, 2001; Knoblauch, 2000) as signifiers acting within a concerted sonic sign system. In other words they are perceived as something else, as standing for something and not just as sounds in themselves. Just as, when we hear a melody we perceive it as melody, rather than as an arbitrary succession of sounds, similarly a person's voice is attended (i.e., interpreted) not merely as acoustic material, but as someone's voice, with its full expressive emotion and its sedimented (textural, embodied, idiosyncratic, social, gendered, cultural, historic) meanings. Our listening is thus a hermeneutic act, we read into sounds and voices and perceive the person as conveying something (a feeling or energetic state) through the voice. For the voice is always heard as alive, dry, sad, happy, vulnerable, close or distant, friendly or hostile – standing in close relation to the uttering person, intimating differing subjective states in the speaker. As we shall see, psychoanalytic theory, or rather, theories, constitute in themselves interpretants through which psychoanalytic meaning is generated.
II.3.4 Peirce’s Sign Classification

Peirce’s fundamental classification of signs into the categories of symbol, icon and index, is the most prevalent in music theory and of great relevance to the conceptualization of the sonic nonlinguistic dimension. Rather than representing objects, each category represents the manner in which the sign relates to the object. Thus, a symbol consists of an arbitrary, unmotivated relation between object and representamen (most words act as symbols). A musical symbol can be considered to be a musical occurrence like a melody, or a motif that refers through convention (musical or other) to an external object - for example, a religious anthem, or a musical motif that conventionally symbolizes a certain feeling in a given culture. For example, the doleful appoggiatura is an expression of grief in Western classical music (Monelle, 1992, p.199). Because of their static, arbitrary nature, symbols are clearly less relevant to musical semioses. Icon and index are the sign types that have the most relevance to music and sound, and will therefore be described in greater detail.

An icon is linked to the object by resemblance: figurative drawings are icons. In terms of music, one could say that music is meaningful because something about it resembles human emotion: the dynamic ebb and flow, tensions and relaxations of human feelings and embodied experience. Langer (1970) and Webb (1760/2003) are representatives of such iconic (or isomorphic) theories, suggesting that music represents the morphology of feelings by analogy. It is important to note that iconicity (resemblance) is not necessarily visual, and that objects can be interpreted

46 Though most of these musical symbols have iconic or indexical roots, as we shall see.
Gina Lucia Brugada

iconically through aural elements. Thus, for example, certain musical motifs are
iconic of sighs, laments, or joyous leaps. Here again, the interpretant of this iconic
signification is the body. It is important to note, that icons often have an element of
stylization, manipulation, and abstraction – thus of convention (road signs, for
example).

Iconicity in sound is an important concept for the analytic sonic dimension. When a
patient talks about a sad event or describes a sad subjective state (or doesn’t, yet
the sadness is made present in nonverbal ways) the sadness is not only told or
described, but is performed, enacted through sonic and gestural icons resembling an
embodied or culturally known trope of sadness that has visual, gestural, as well as
sonic typical traits: for example, drooping facial features, a slowing of gestural and
vocal pace, a low or shaky tone of voice, etc. Sonic icons importantly involve the
analytic listener’s conventionalized expectations. Iconicity in nonlinguistic sound
helps listeners identify emotional themes not just cognitively: recognition of affect
through sonic and gestural levels occurs through identificatory, embodied processes,
making interaction into an empathic, intersubjective participatory event.47 This
interpersonal process is considered by many infant researchers as essential in the
development of intersubjective (Stern [1985] for example coined it as “affect
attunement”, a concept that will be delved into in the next chapters).

It is important to consider the interaction between the concept of musical icon and
psychoanalytic thought. The use of musical-semiotic formulations in the analytic

\[47\text{Recent literature on mirror neurons supports this direction of thought. Vickohff and Malgrem (2004) maintain that mirror neurons are one of the explanation to the question why does music move us.}\]
situation, suggests, to some extent, the treatment of vocal productions and listening acts as manipulated, stylized and even composed in a musical sense. My argument is that there is a certain measure of intention in a person's use of voice and response to voice and other sonic elements. As has been argued in music theory and literary theory, voice is, in many ways, an effect of individual and social forces, reflecting values, beliefs and practices. Within the psychoanalytic arena, it is suggested that the individual, social and historical markers reflected in the voice can be related to the unconscious. If this is the case, voice can be considered to act as a symptom (and then the semiotic assumption is that voice and unconscious states act in some way as sign and signified). Thus the voice, along with its attending components, acts as symptom not only in an expressive capacity as is commonly assumed. Voice both expresses and articulates (if voice would only be expression then there would be a danger of positing a fallacious direct line between voice and self). For example, the voice expresses sadness, but at the same time it articulates the sadness through vocal timbre, gesture and rhythm while stylizing it in a certain way. Of course, manipulation is here not meant in a cynical sense (although sometimes that is certainly the case), but rather as an unreflective semiotic articulation.

---

48 There is a danger of thinking that sonic signs point to a reified interiority. Sonic signs, like linguistic signs, are here conceived to generate, through their utterance, a subjectivity. Nevertheless, a first-person experience is postulated, only a part of which is visible or audible (perceived from the outside). From this point of view, as suggested by Cumming (2000), the self is articulated through various inner and outer manifestations that work as signs in a larger structure, part of which faces inwards (only the individual experiences it), and part - outward, facing out towards the other. However, later on we shall see that also this distinction is problematic: Bakhtin (as Lacan), for example, posits the existence of otherness in the self, furthermore, in the actual possibility to conceive of the self. As we have seen above, Merleau-Ponty's views also counter the binary division of inner and outer.
Psychoanalytic interpretations of these vocal signs will be critically reviewed in the next chapter.

To resume the theme of sonic icons: A patient’s or analyst’s speech, with its genre, stylization and idiosyncrasies can act as an icon, aurally resembling certain psychic and affective configurations. As relatively stable categories of sonic expression, icons can be taken to be conventionalized sonic images or configurations that can act as vehicles for the performance of more individualized, nuanced expression of particular, idiosyncratic idioms of affect (the expression of sadness is always an interplay between a culturally established style of expressing sadness and an individual qualification of its expression. Thus, the individual can be viewed as manipulating – albeit prereflectively or unconsciously - sonic and gestural stylistic elements in expressing an individual nuance of sadness). In this study I propose that sonic nonlinguistic signs are not, as is commonly believed, simply “natural”, nor are they originary, direct expressions of speakers’ states. Although vocal nonlinguistic expression is indeed more direct and less mediated than verbal speech (i.e., voice reflects affective states synecdochically or indexically – I will come to this in a moment), nonetheless I am suggesting that it reflects a certain- even minimal - amount of manipulation.\textsuperscript{49} The idea of iconicity suggests that the sonic dimension is a construction containing style, genre, historic and social situatedness. In creating certain affective and energetic icons, as constructed individually by each analytic participant and by the dyad, voices in analysis can be viewed, to some extent, as “staged” and stylized. The dialogue found in speech between conventionalized sonic

\textsuperscript{49} See Dolar (2007), about the poetics of coughing: even in such physically natural manifestations such as coughing, one could discern a semiotic activity.
icons (sonic dialects) and individualized expressions (sonic idiolects) is an important area of investigation. Voice, when taken as performance in the analytic scenario, is relevant to the issue of the unconscious because it posits the unconscious dimension of sound or voice as symptomatic: through its articulation, sound as a malleable medium is capable of being manipulated and cathexed, i.e., unconsciously charged with meaning through performative acts. The performance of the sonic dimension in the clinical context can thus be viewed as a form of sonic acting out.

The third category of sign, the index, is according to Peirce a type of sign that is directly affected by the object and is contiguous to or a part of it. According to Cumming (2000), nonverbal indices suggest a factual situation – smoke suggests the existence of fire and a cry suggests the state of pain. Coker (in Monelle, 1992) understands the index in the same modified sense. An index is a pointer, it focuses one’s attention by attracting attention and then directing it to an object, specifying more or less the location or state of the object or event in space or time. An example of this sense of the index is a pointing gesture, an arrow or the sound of an approaching train. In Western music theory, the indexicality of music is considered to have its origin in the intonation of the voice. The intonation and coloring of the human voice are the most salient forms of indexicality as they are directly connected to embodied aspects of affect and energy. According to Monelle (ibid.), the signifying function of this sort of dynamic index is synergistic, based not merely on contiguity but on contiguity in motion. Thus, music and emotions share a natural

---

50This idea has been advanced in the fields of sociology and anthropology as well as in performance theory (see Schechner, 1988 and Turner, 1986). Several analytic authors also engage the analytic encounter as ritual or staged performance (for example, Ringstrom, 2001).

51In fact, what psychoanalysis views as symptoms can be considered, in a non pathomorphic manner, to act as indexical signs.
order (Webb, 2003/1760), rendering the relation of music and feeling metonymic rather than metaphoric, indexical rather than iconic. Similar contentions about the affinities between sound and affective life – the roots of intersubjective life – are held in developmental theory (Stern, 1985; Malloch and Trevarthen, 2009; Trevarthen, 1999-2000) and in music therapy theory (for example Bruscia, 1987, 1989; De Backer, 2004; Dissanayake, 2001; Nordoff and Robbins, 1977; Pavlicevic, 1990, 1997; Sekeles, 1995, 1996; 1996a; Streeter, 1999a, 1999b). Sonic phenomena often act as indexes to affective and bodily states - vocal inflection, rhythm, and pause point the listener's attention to these states or occurrences in the speaker. In turn, the awareness of these states acts as indices to other referents, like internalized relational patterns, traumatic experience, here-and-now reactions, etc. Vocal tone in psychoanalysis is often seen as a channel for expression of affects and ideas, that, intentionally or not, cannot be verbalized. Voice is implicitly understood as indexical (pointing to and part of) the speaker's state (whether patient or analyst). The various forms of slippage and discrepancy between verbal and vocal nonverbal messages provide an important area of inquiry. Thus, as described by Stern (1985), while word is considered to be officially accountable for (it is the individual's “official” account, thus is undeniable), vocal nonverbal expressions, because gradient and unrepeatable are deniable. On the other hand, these expressions are paradoxically read as the ones that were really meant, that is, more authentic than verbal messages. Thus word and nonlinguistic voice embody an interplay between symbol and index, creating complex significative effects such as irony, double-bind messages, authenticity and falseness.
The distinction between these three types of signs is, of course, an abstraction, for
often times signs contain in them several signifying functions- either symbolic, iconic,
or indexical or any combination of the three. Cumming (1996, 2000), for example,
discusses the embeddedness of sonic indexes in iconic musical signs (or, how
indexes can transform into icons). This consideration is important in explaining how
local sonic events create more abstracted signifying structures. For example, a
patient’s sudden sigh, imperatively causes an analytic listener to direct attention to the source of the speaker’s distress, thus indexing a certain local affective state. At
the same time sighing presents an aural icon of distress and in less acute
circumstances of vocal production (for example when sighing is a recurring motif in
one’s speech) though the indexing of a physical state in the voice remains evident, it
ceases to be of imperative effect in itself and acts iconically as representing a certain,
more stable affective state (for example, of a stance of concern or worry).

The relevance of Peirce’s concepts of signs to psychoanalytic practice is obvious.
Analytic activity consists of interpreting surface events (whether linguistic, nonverbal
and relational) as signs standing for psychic meanings. Meaning is here considered
to generate from the semiotic (linguistic and nonverbal) interactive activity of the
participants (see Canestri, 2000, Killingmo, 1990, Kristeva, 1986a, 1986b, and other
analytic writers with a semiotic orientation). Peirce’s sign typology permits us to
think of analytic clinical discourse as a semiosis composed of different kinds of signs
and levels of signification. Some are more transparent and arbitrary, others – like
indices - are more implied and motivated. Whereas psychoanalysis dealt mostly with

52 Peirce commented on indices that “direct the attention to their objects by blind compulsion” (in
Monelle, 1992, p.212)
the linguistic sign (Peirce’s symbol), the concepts of icon and index, that are not propositional (as they do not assert anything) make it possible to articulate more elusive semiotic activity. The view of analytic semiotic activity as an interaction between symbols, icons and indices, displays a richer and more complex picture, making it possible to attend to the semiosis of nonlinguistic sound in speech simultaneously with sound-as-word, in several capacities and varying forms. Similar ideas, though formulated differently, have been exposed by various authors from cognitive- and neuro-science (for example Bucci, 1985, 2001, Pally, 1996, 2001; Schore, 1994), psychoanalysis (Killingmo, 1990, among others) and semiotics (Kristeva, ibid.).

**II.3.5 Semantic, Rhetoric and Narrative Processes**

By articulating certain sonic categories, like sonic gesture, posture and mood, I have dealt with the material side of the sonic dimension in semiosis. But, as I claimed above, and in accordance with Peirce's idea of the sign, meaning is not found in the ontology of sonic signs, but rather in processes of signification that lie in the interaction between signs, objects and interpretants in several levels of occurrence (verbal meaning, gestural signs, etc.). Therefore, the second focus of this inquiry addresses the question of how these sonic categories work. How do they deploy in time and how are they bound together in creating larger-scale sonic structures and effects? For it must be remembered that sonic signs are temporal, processual signs. Further categories thus will be designated to explicate the semantic, narrative and rhetoric operation of sonic signs in their continuous emergence and unfolding. In
this capacity I shall investigate, in chapter III and IV, several sonic phenomena and signifying mechanisms, inspired by music theory.

Repetition, for example, is a central device in music and nonverbal sonic fields. Lidov (2005) and Cumming (2000) elucidate the phenomenon of repetition in music with its several syntactic and semantic capacities. They demonstrate that repetition, in its manifold variations, can create a number of phenomena with rhetoric, semantic and narrative effects on local and structural levels. In the context of the analytic exchange, it will be demonstrated how vocal repetition can create varying effects such as rhetoric of affirmation, increase of intensity, or effects of stasis or circularity. The relation between this and the important psychoanalytic outlook on repetition will be elucidated.

Another sonic device that will be discussed is the employment of sonic invariants to create norms and expectations on the one hand, and surprise and innovation on the other. This is brought about by the successive flow of sonic events that create effects of sameness, contrast and opposition. Through the concatenation and nuanced combination of sonic elements, quasi-narrative effects of an affective nature are created. The stream of dynamic unfolding sonic events creates a build-up of tensions, resolutions, emergences and dissolutions that are not only highly evocative of embodied and affective narratives but enables the creation of sonic metaphors of psychic time, space and relations. These were addressed in the field of music by authors as Abbate (1991); Almén (2008), HaCohen (1999, 2001b), Langer (1970), Tolbert (2001, 2002) and Zuckerkandl (1957). At the same time, meaning results from the effects of the simultaneity (and not only succession) of sonic and
other phenomena, creating more stable, structural effects and categories such as sonic persona (a sonic “character” of the musical work or the performer) and sonic ambience and mood. This continually unfolding dynamic, “polyphonic” texture intersects affective and propositional, individual and social, material and processual synchronic and diachronic levels of signification.

II.3.6 The Dialogic Interaction of Signs

It will be noted from the above that Peirce’s model of semiotics highlights semiosis as activity, i.e. as processes of relating objects to representamens. These dynamic processes are conceptualized in this research as dialogic activity between several internal, external, local and structural levels of signification. Hence, a sign can be interpreted through its internal dialogics: as we have seen, voice demonstrates an interplay between structural binary elements (symbolic functions) and analogue, presentational elements (indexical or iconic functions). In addition, the dialogical nature of semiosis can be considered as a dialogic interplay between local levels of the uttering individual (through phenomena as vocal timbre and gesture in the speaker); more complex phenomena occurring in the dyadic level (sonic phenomena occurring between the two participants, such as vocal duet), and ultimately globally (through categories like vocal mood that although created by the analytic dyad, reflects a distinct entity that unfolds beyond the individual or the bi-directionality of the dialogic exchange\textsuperscript{53}).

\textsuperscript{53} As I will claim in chapter III, vocal ambience is here seen as a manifestation in sound of Ogden’s concept of the analytic intersubjective third (1999b): Vocal atmosphere, though created by the analytic partners, can be heard as a third subjectivity, that envelops them, as it were, and influences their interaction.
This fundamentally dialogic conception underlies the sonic nonlinguistic dimension and is based on Bakhtin’s theory of dialogism, the third theoretical perspective to which I now turn.

II.4 Dialogism and Sound

Until this point I introduced phenomenology and music theory as two theoretic templates that facilitate the apprehension of sound from an inner phenomenological, first-person experience, and an outer semiotic viewpoint that makes possible the creation of categories and articulations. These two positions focus on the materiality of sonic phenomena, their deployment and processes of signification. Now I turn to a third perspective, whose unit of reference is the dialogue. The theoretical base for this approach is the literary critic Mikhail Bakhtin’s concept of dialogue. It must be clarified that this dialogic focus is not used as yet another aspect of language and of sound’s meaning. The dialogic perspective I employ here, following Bakhtin, is more total and radical: dialogue is the central template in the understanding of communication and any occurrence (in our particular case, sonic utterance) in the analytic encounter can only be understood as being generated from a dialogic matrix. Thus relation, in a Bakhtinian sense, takes center stage, constituting a key notion for understanding the sonic realm (note, however, that this is not the psychoanalytic take on relation, i.e. unconscious forces like projection, transference and countertransference, though these important concepts will be engaged later on).  

54 Relation and dialogue in Bakhtin’s universe

---

54 However, relational psychoanalysis has recognized the affinity of Bakhtin’s theories to their basic tenets, as evidenced by writers such as Beebe et al. (2003b).
have special meanings, and before I proceed to talk about sound as a dialogic occurrence, Bakhtin’s concept of dialogue must now be explained.

**II.4.1 The Significance of Bakhtin’s Dialogue**

Throughout his work, Bakhtin uses the idea of dialogue in several senses, essentially in relation to the self, language and literature (Morson and Emerson, 1990).

Dialogue is a complex of ideas that refer to human behavior as reflected by the use of language. Since the present scope does not permit to discuss these ideas fully, I will in the following section sketch out, however briefly, a more general sense of the meaning of Bakhtin's dialogue, after which I will expound certain facets thereof that are relevant to the present discussion.

In a general sense dialogism\(^{55}\) refers to a global and existential level: it is used as a model of the world; an epistemology that explains in dialogic terms the generation of meaning in the human realm, from the most local level of dialogue between two individuals, to relations between social groups, ideologies, political bodies, cultural entities, and art genres. Thus Bakhtin states:

> Dialogic relationships are a much broader phenomenon than mere rejoinders in a dialogue laid out compositionally in a text; they are an almost universal phenomenon, permeating all human speech and all relationships and manifestations of human life – in general, everything that has meaning and significance. (Bakhtin, 1963/1984, p. 40)

---

\(^{55}\) Bakhtin himself didn’t use the term “dialogism”. It was coined by his eminent interpreter and translator Michael Holquist (1990) as an encompassing category for the various meanings of dialogue in Bakhtin’s universe.
Bakhtin’s dialogue is pragmatically oriented. He vehemently critiques “theoretism” (a concept he coined signifying his denigration of rationalistic trends towards theoretical abstraction). For Bakhtin, philosophical trends that subsume the unfinalized, “messy” qualities of real life, communicative and social acts under a comprehensive, closed explanatory system end up in suppressing, and ultimately losing, the vitality of the everyday social world and its concrete particularity.

Bakhtin’s dialogue differs from Buber’s I-Thou relationship and is in stark opposition with Hegelian and Marxist dialectics. As we shall see, Bakhtin’s dialogue is never neatly unified or synthesized. More than anything, it constitutes an event - the event of existence - that is always new, always of blurred edges and a messy (or as nonlinear systems theory would have it – “sloppy”), unfinalized nature, where utterances are fundamentally unique and unrepeatable. It is these qualities that make human exchange and language vital, creative, instilling a sense of hope, agency and an open future for the individual. In this sense, Bakhtin shares some of Bergson’s essential views on the human being (1912/2007), especially the notion that the human relation to life is one of interested, purposeful activity, choice and participation.

In this Bakhtin stands in strong opposition not only to structural linguistics as represented by Saussure, but also to major traditional trends that were concerned with language: stylistics, rhetoric, semiotics and poetics. These traditions, contended Bakhtin, were flawed in their basic suppositions about language and resulted in a deadly reification of language as an abstract structure.
Most importantly to the present study, Bakhtin objected to the Saussurean division of language into *langue* and *parole*, arguing that, although such binary splitting may be useful for analytic purposes, it leads to a misconception of the nature of utterance (Bakhtin's central concept, that will be described below). The fundamental position of the present research is based on precisely this criticism of the binarism of structural linguistics and its lethal consequences for the voice.

Bakhtin tries to elude the fixity of words. In his writing he gives rise to another kind of speech about speech, voice and dialogue, emphasizing the open, dynamic, ever-evolving and indeterminate nature of dialogue. At times he uses neologisms like “heteroglossia”, or “dialogization”, and addresses human utterance through a rich use of metaphors like “polyphony”, “authoring” and “event” creating a new language that depicts his personal vision on dialogue, in fact, *performing* dialogism in his own poiesis. In the language he invents he tries to generate a realm where words produce “a living and unrepeatable play of colors and light” (1981, p.277), “disperse into rivulets and droplets” (ibid., p.263) or reverberate in “harmony and disharmony…dissonance and consonance… [in] a positive project of creating a style” (ibid., p.283). These metaphoric phrasings clearly intimate an affinity between Bakhtin’s vision of dialogue and musical experience.

Inherent in dialogue is a sense of action and vitality that Bakhtin called “the eventness of being”. The idea of event is a complex one, containing a combination of several concepts. It can be understood to be the dialogic process occurring at each moment between the uniqueness of the uttering individual (from its unique place and time in the world) and the multiple discourses coming to him from the outside
world. In this sense, existence as event is dialogic, a “co-being” of self and other - the paradoxical state of needing others' views and categories to define and express the self's uniqueness and creativity (Holquist, 1990). Bakhtin’s eventness is reminiscent of Winnicott’s concept of the paradox inherent in transitional phenomena and the potential space (1971), both created in the individual’s psychic space that exists at the meeting point between outside reality (or otherness), and the inner world. The eventness of existence (the daily event of being) is characterized by a sense of action, particularity, unfinalizability, creativity and openness to the future, representing unique individual choices among the infinite possibilities coming to us constantly from the world. Most importantly for psychoanalytic thought, the eventness of being entails a sense of ethical responsibility towards the self and the other. Thus events (dialogic activity) are to be understood as responsible choices the individual must make in the course of existence.

Later on we shall see the centrality of sound in the constitution of speech-events.

**II.4.2 Utterance**

I now come into a second sense of dialogue that is even more pertinent to my present concern and must be delved into more thoroughly. Dialogue is about what happens when two people meet through their use of language. However, dialogic interaction for Bakhtin is more complex than meets the eye and cannot be reduced to simple verbal interaction. In order to grasp its meaning, the concept of “utterance” must be introduced.

---

56 In this respect there are important parallels to be drawn between Bakhtin and Levinas, as illuminated by Erdinast-Vulcan (2008).
Utterance can be seen as the basic unit of investigation, or building block, in the Bakhtinian dialogue. Utterance is the act of saying something and contains within it a voice in the act of speaking.\(^{57}\) To be an utterance, words must be acted out, performed. Utterances are the fabric of live speech, converting each dialogue into a unique, singular event. Bakhtin differentiates utterance from sentence or word, and, as we have seen, this differentiation is crucial in its departure from traditional linguistic views of language. An utterance is not equal to a sentence: it can be a nonlinguistic gesture (like a sigh), a word, a sentence, or a whole speech. Later on we will see how utterance is closely related to sound and the voice.

However, though the utterance is central in expressing the self, it always contains an implicit dialogue between self and other. The self's utterances are not to be mistaken for a private, “inner”, originary communication, since the subject is not a monadic entity transmitting apriori messages to a listener. As we have seen, the Bakhtinian subject is permeated throughout by alterity, as historical and social forces are sedimented in self-perception, language and speech. The necessity to use language exemplifies the inevitable embeddedness of the self in external factors. Still, for Bakhtin (in opposition to the majority of structuralist and deconstructionist views) the subject is not totally passive, a mere effect of social forces, but is uniquely situated in a center – characterized by a time and a place - from where she must actively respond and dialogue with these external forces. To restate, Bakhtin is not positing a privileged ontological center for the speaking subject. On the contrary, the other too, when considered from a first-person viewpoint, also operates from his own center, his own irreducible perspective. However for the subject's existence to

\(^{57}\) The concept of utterance resonates with Austin's (1962) speech-act theory.
have meaning, he needs the other as a defining background and must impose the other's categories, to stabilize and objectify to a degree, the indeterminate flux of his own existence. Through intersecting the other's words and categories with her own, personal meanings (coming from her own personal position and history), through her unique performance and response, the subject constitutes her distinct individuality. Hence, every utterance contains an internal dialogue - the ongoing negotiation between the individual speaker's need to communicate something and the global constraints of language (representing otherness, history and society) as an external generalizing system. From a sonic point of view, it is important to see that these dialogic relations exist within the sonic dimension as well. Thus in a single word there is an internal dialogue between sonic components used for phonemic articulation and those components used more idiosyncratically in the course of expression. The phonemic sounds represent stable, repeatable, shared aspects of a language (centripetal or unifying forces of a language), while the expressive, qualifying sounds represent individual, unique, unrepeatable aspects of speech (centrifugal or dispersing forces in a language).

The multi-layered nature of language is called “heteroglossia” by Bakhtin due to the multitude of social dialects, jargons, identities, roles and socio-ideological aspects that accumulate in language over time and are sedimented in speech. Language is not a neutral medium that can be simply seized by a speaker, but something that comes to us populated with the history and intentions of others. Every word is
charged with the contexts in which it has lived, whether social, historical, or uniquely personal (it is the latter that often occupies the psychoanalytic process).58

Dialogism thus claims a multiplicity and simultaneity of perspectives that is not dialectical or reducible to a unifying formula, but exists as an open, unfinalizable dialogue between self and other. Closely related to the idea of multiplicity of utterances and discourse is the idea of polyphony, which, beyond its obvious connotation to the idea of a multiplicity of voices, implies, according to Bakhtin a unifying consciousness – a super-consciousness that coordinates, or dialogizes these different voices. But in the dialogic conception this unifying consciousness is not a totalitarian one. It represents at the same time another voice, and is affected by the individuality of the voices it is trying to unify. Thus writers and composers become equal to the characters or voices they are trying to create, in a sense being created by them. Similarly, an individual, in the act of authoring himself (more on this in chapter III), can represent simultaneously a multi-vocality, or heterophony, whereby different voices and selves are dialogized.

II.4.3 Addressivity

As we have seen, in dialogism every utterance always exists in response to something that has been said before. But at the same time the utterance turns outwards, towards an interlocutor, inviting a response and yet anticipating it. Thus,

58 The individual’s dialogue with existence is seen by Bakhtin as a continual process of creating – or, as Bakhtin phrased it – authoring the self. Listening to the voice of the speaker is listening to the subject in the act of authoring himself. The concept of self as author will be related, in the next chapter, to the concept of vocal persona which I presented earlier. Here I will briefly say that the vocal persona is the subject in the act of authoring herself sonically.
words are never spoken in a vacuum and do not operate as abstract, independent entities conforming to a logical, stable systems of laws. Speech (and the language it uses and the ideas it communicates) is dynamic and relational. 

Utterances are therefore motivated by a structure of address, and cannot work as utterances without an addressee to whom they are directed. Bakhtin coined the concept of “addressivity”\(^{59}\) to indicate that utterances have a certain inner energy, an impulse towards someone, the listener. Addressivity is inherent in the utterance, not in the words, and in order for words to become utterances, they have to emerge from an impulse towards the other.\(^{60}\) In his seminal essay Discourse in the Novel (1981), Bakhtin states that

\[
discourse lives, as it were, beyond itself, in a living impulse toward the object;
\]

if we detach ourselves completely from this impulse, all we have left is the naked corpse of the word, from which we can learn nothing at all about the social situation or the fate of a given word in life. (Bakhtin, 1981, p. 292, my italics).

Bakhtin’s concept of addressivity can also be understood through dialogue’s tripartite basic schema, that features three basic elements: an utterance, a reply, and the relation between the two (Holquist, 1990). This schema interestingly resonates with Peirce’s conception of the sign as shown above - the tripartite construct of object, sign and interpretant. It is clear that for both Peirce and Bakhtin,  

\(^{59}\) Addressivity is sometimes conceived by Bakhtin as “answerability” or responsibility. Here again, as in the concept of event, addressivity bestows the dialogue ethical overtones. We shall later see how answerability and responsibility resonate in the voice, for example in the sounds of empathy.  

\(^{60}\) Later on we shall see that within the musical sphere, sounds and sonic configurations are also understood and interpreted by listeners through their impulse and directionality to other sounds (Zuckerkandl, 1957).
the most important element in their schemes was the third and most abstract
element, the process that connects between two elements (Peirce's interpretant
and Bakhtin's relation).\textsuperscript{61} In other words, both Bakhtin and Peirce place a central
emphasis on a human dynamic happening-- an intention, impulse and process – in
short, addressivity - that connects two elements by generating context-specific, non
pre-determined meanings. This conception provides a special means to
understanding processes of signification, and proves to be important to the
understanding of the nondiscursive sonic realm. Instead of viewing the voice as a
reified separate entity that can be seized and formally analyzed (a linguistic
perspective), notions such as dialogue and the interpretant enable us to conceive of
communication as relational energy, impulse, intention and gesture.

\textbf{II.4.4 Dialogism and the Sonic Dimension of the Psychoanalytic Encounter}

The above ideas have a special relevance to psychoanalysis, because in the
psychoanalytic situation, live speech is central.\textsuperscript{62} As we know, communication here
is not merely the transmission of verbal messages aiming at intellectual
understanding and interpretation. Psychoanalysis can be seen as a purposeful
staging of an interpersonal encounter, where communication processes (dialogue)
are intensely scrutinized. Over the past hundred years psychoanalysis has carefully
elaborated the dialogue in the clinical situation - indeed a singular form of dialogue,
where, through live speech and a particular form of listening, curative changes are
achieved in the life of the subject. The analyst works from a specialized listening

\textsuperscript{61} The affinity between Bakhtin and Pierce is also acknowledged by Holquist.

\textsuperscript{62} Several contemporary psychoanalytic authors have recognized the importance of Bakhtin's
dialogism to the psychoanalytic project (for example Canestri, 2000 and Priel, 1999).
stance, attending to the transference relations that emerge through the dialogue. In Bakhtinian terminology we can say that the analyst positions himself predominantly as addressee, and that his experience *qua* addressee of a particular patient sheds light on the rich dialogic activity that goes on in the patient’s mind. From this perspective, both dialogism and psychoanalysis understand speech as something that is beyond its immediate, literal meaning. But, as I will further explicate in chapter III, psychoanalytic listening is carried out primarily from a linguistic vantage point and attends mostly to the linguistic (denotative and connotative) level of speech. Because it originates from a classical propositional paradigm, psychoanalysis could not conceptually encompass the operation of the voice and the pragmatics of speech (for an account on the linguistic paradigm affecting psychoanalysis see Canestri, 2000).

How do concepts such as utterance, event, addressivity and polyphony help in understanding the sonic dimension in psychoanalysis? First - they enable the listener to interpret what is said not merely linguistically through the reductive semantic-logical-referential paradigm of the logos; but, after Bakhtin, through a dialogical template. Accordingly, the listener takes in the word not just in its denotative and connotative literal meanings, but rather, he takes in the utterance complete with its

---

63 Clearly, within a dialogic conception, this is a mutual happening, and the patient acts also in the capacity of the analyst’s addressee. Relational psychoanalysis acknowledges this mutuality though still asserts the asymmetry in patient and analyst roles.

64 But while Bakhtin addressed the individual through his relations to social groups, Freud viewed the family romance as the central influence of the individual psychic life, which he consequently extended (projected) onto social experience (Holquist, 1990). As a consequence, psychoanalysis is attentive to the effect of *personal* (historic, intersubjective) experiences as they are reflected in language. Another crucial difference between the two is that Bakhtin was suspicious of the Freudian notion of the unconscious - as he was of all deterministic or causal explanations of the human being. Bakhtin was inclined to explain human existence through a richer, more varied concept of consciousness (Morson and Emerson, 1990).
voice and gesture, its emotional impulse and directionality towards another - its 
music. To hear the individual's performed speech is to hear the emotional, 
intersubjective world that lives in the speech. Every utterance – whether linguistic, 
as a word or sentence, or nonlinguistic, as a pause or an accelerated tempo – inheres 
in the dialogic context or strictly speaking, in a specific structure of addressivity. No 
utterance (word, gesture, voice inflection, silence, switching pause) stands 
automonomously or has an originary source. It is always situated in a communicative 
web, and its meaning comes about from its existence in a dialogical field. The 
intimate speech, the bitter speech and the dead speech in psychoanalysis are 
constituted beyond their possible discursive meanings; these utterances unfold with 
their characteristic music through the sounds and rhythms of intimacy, bitterness, 
and deadness – alerting the listener to introjected dialogues with past others.

Bakhtin alerts us to the fact that the same sentence will mean something completely 
different when it is said intimately, bitterly or in a flat, unexciting way. His concept 
of dialogue provides a way to listen to these sentences as live utterances, as “songs 
of experience” where voice, rhythm and words co-create meaning.

From a Peircean semiotic point of view, the concept of addressivity works here as an 
interpretant in the process of listening. When a patient says something sadly, his 
utterance is interpreted not exclusively by isolating verbal content from her sad 
voice, nor through a sonic formal analysis (how is sadness shaped by the voice), but 
through the idea that it is an utterance directed at someone (concrete and/or 
imagined listener), while simultaneously responding to a previous utterance. Thus, 
content and form are configured by the structure of addressivity: in the most
fundamental manner the sad utterance is addressed to me as listener (though I may be representing and including other listeners). If the patient addressed her mother or best friend, the utterance would be different (perhaps its sonic shape and “feel” would be different and other facets of its content would be communicated; for example, sounding more angry-sad than desolate-sad). Thus, on the first, immediate level, the patient’s utterances are directed at the analyst, which means that words, narratives, behaviours and speech acts are meaningful not only in themselves (in their content), but they represent the choice of content and form of the patient’s narrative and presentation that are evoked vis-à-vis the specific person of the analyst. The patient’s utterances are not only directed at the analyst, but are precipitated by the person of the analyst. Similarly, the analyst’s utterances are precipitated by the patient’s presence and performance, tone and manner. The interlocutor, (the addressee) is in fact one of the determinants of the utterance. Phrased in a reverse way, the listener is an active co-creator of the utterance (more of this will be said in Chapter IV).

Secondly, the dialogics of Bakhtin makes it possible to address the totality of living speech that, beyond its semantics, resounds in the intersubjective field: The dynamics of interaction and interpersonal space as reverberating through sonic texture (volume, tone, atmosphere, gesture, norms, expectations). Dialogism makes audible the “sound of relations” or relations as sonic events in the process of becoming, unfolding or dissolving; in their arrival at intersections and unstable locations; in moments of movement and change; in response and reaction; in anticipation and annunciation, in the emergence of articulation and transformation;
at switching points, stasis and immobility as well as in moments of breaking through, surprise and excitement. Dialogism in other words, amplifies the soundings of utterance as a semiosis of process, transaction, intentional reaching out, convergence, divergence, attunement, misattunement and co-construction. As we shall see, these semioses crystallize into recognizable sonic relational forms, evoking musical phenomena such as harmony, dissonance, fugue (chase), repetition (as in rondo or theme and variations forms), call and response, and others.

Thirdly, it can be said that the dialogic perspective makes compellingly present other sounds, rhythms and gestures than those occurring at a given time between interlocutors. The sonic dimension of speech implicitly echoes more distant dialogic actors. The present concrete dialogue is interlaced with dialogues with the past voices of significant others in the history of the participants. For the voice in dialogue, by making those relations audible, makes the sounds of the past audible in the present. The interrelations between present and past dialogic sounds and rhythms create a veritable polyphony where implicit norms, values, expectations and beliefs are voiced in individual utterances. Dialogism sensitizes the listener to hearing the emergence of meaning through speech-sound in a multidimensional way.

II.5 Theoretical Interrelations

Being an interdisciplinary project, one of the goals of the present research is the dialogization of conceptual worlds dealing with different objects and phenomena in
the human sphere. Such dialogization endeavors to juxtapose and negotiate different viewpoints on similar phenomena. In this manner, phenomena can be understood and recast through differing points of view and understandings that exist simultaneously, thus enriching their possible signification. Throughout the exposition of phenomenology, music semiotics and dialogism, I have hinted at their certain interrelationships. Now I shall more clearly indicate some common themes and concerns between the three. It is not my intention to unify these paradigms or subsume them in any way under a common idea, but to show commonalities and meeting points that create a shared ground through which sonic phenomena in speech can be more richly conceptualized.

II.5.1. The Centrality of the Body.

In this study, the sonic dimension as existing in communication is understood primarily as an embodied phenomenon. In all three theoretical perspectives, though conceptualized differently, the body plays a central role. Basically, the body (as body-self) is the fundamental site from where acts of perception, consciousness, action and interaction are generated. While phenomenology posits the body as a fundamental locus of meaning, music-semiotic concepts such as vocal timbre, vocal gesture and vocal persona emphasize the body as a primary object of sound-signs. These act as body indexes and icons, and from a phenomenological point of view it can be claimed that a first-person bodily perception and consciousness is at work in the attribution of meaning to sonic events. Bakhtin's concepts of dialogue, utterance and event are also based on the immediacy of lived experience of the self as a body-self. His main philosophic concerns are derived from the body's
boundaries and limitations, its location in a given time and space, and its unique point of view on other bodies and on the world. Although Bakhtin’s dialogism is not a phenomenological philosophy, it shares with it an advocacy of the individual’s point of view as a center of experience.

Also common to the three perspectives are their theoretical constructs that address human communication through other paradigms than the binary paradigms based on structural linguistics. In fact, they reject, sometimes with fervour, these paradigms and claim that they are inadequate in explaining the richness and vitality of human symbolic activity. For our purposes, nonlinguistic communication that is bodily performed through gesture, vocal tone and rhythm and their subtle interactive coordination must be explicated not as complementary addendums to verbal language, but through wider, more holistic paradigms that repair the traditional rupture between psyche and soma.

II.5.2. Contextualism

As we have seen, phenomenology, music semiotics and dialogism share a certain conception about objects and their relationship to their surroundings. Accordingly, phenomena (external like objects, or internal like perceptions and feelings) are never apprehended as isolated entities and can only be understood against the background of their surrounding environments. It is the background that gives the object its meaning and can be seen, in fact, as an active element that is (paradoxically) more decisive than the object in determining its meaning. Backgrounds consist of anything from concrete, physical contexts where phenomena are located, to human backgrounds (significant others, social groups), historical situatedness (personal and
social history and ideologies as context), and semiotic context. Sonic signs, events and practices must be read in a temporal-spatial (or historical) situated context.

Contextualism is thus a common theme in the three theoretical fields and a leading principle in the present research.

Closely related is the idea of first-person experience as a fundamental, though shifting, center of reference – an idea featuring in some form or another in the three theories. In all of them, meaning is generated in a multidimensional space and time created by a dialogization of multiple, shifting centers of meaning, through oscillating (nonlinear) attention and perception. Oscillating attention is here viewed as an all important possibility - as a dialogizing act and stance that indicates the listener's recognition and responsibility towards differing points of view and the commitment to an encompassing perception – the polyphony of multiple discourses.

However, it should be stressed that this is far from being a relativist stance. Although the centrality of language (and thus, the impact of otherness in the subject's predicament) is acknowledged by the three theories, they posit language not as a neutral, abstract entity, but as representing the subject's active interaction with the world in the struggle to create an individual voice. Bakhtin states:

... language, for the individual consciousness, lies on the borderline between oneself and the other. The word in language is half someone else's. It becomes 'one's own' only when the speaker populates it with his own intention, his own accent, when he appropriates the word, adapting it to his own semantic and expressive intention. Prior to this moment of appropriation, the word does not exist in a neutral and impersonal
language... but rather it exists in other people’s mouths, in other people's contexts, serving other people's intentions: it is from there that one must take the word, and make it one's own. (1981, pp. 293-294)

For, contrary to relativistic and deconstructive views, in the midst of contextuality and the multiplicity of meaning, there exists an opportunity for the subject to create a unique voice, leading to a possibility of truth and authenticity. Through recognition and dialogue with the multiplicity and unfinalizability of existence and the concepts of dialogue, body-subject and the semiosis of the voice, the possibility of the individual's sense of creativity and choice is restored. The present study investigates the ways and means words become “one’s own” through sound and how the realm of sound works in the creation of one's own voice.

**II.5.3 Polyphony: A Unity of Multiplicities**

In juxtaposition with the topics of multiplicity and contextuality, the three theoretical approaches share certain unifying forces, or a unifying consciousness that dialogizes different voices and perspectives. Thus, various glosses are put in interaction to create a complex, multidimensional space, rich in meaning and perspective. Yet, the task of the unifying consciousness is not to provide for a final, clear-cut solution. In dialogue there always is an unfinalizability, an open horizon of meanings and the upper consciousness does not control the different voices, but is merely one of them. This point evokes several phenomena such as polyphony,

---

65 Similar conceptions are advanced in psychoanalysis by authors and clinicians whose conception of the self is more of multiple selves who are in different kinds of relations (see for example Bromberg, 1994). Dissociation, for example, is explained by Bromberg as a defense mechanism where disconnection between two selves or self experiences is effected to avoid psychic pain. Psychological improvement is considered thus as the ability of the individual to recognize and negotiate between his different selves or experiences, or, in Bakhtinian terms, to dialogize his different self experiences.
heterophony, jazz improvisation, dialogue and heteroglossia. Coming from the fields of philosophy, literature and music, these phenomena represent multi-vocal events where languages, sounds or utterances (each operating from its own center and consciousness) coalesce into polyphonic textures.

**II.5.4 How does it work?**

To exemplify in what manner the three approaches interact, I now offer as an example a hypothetical sonic experience in the clinical arena. While in the following I exemplify the input of these theories separately, needless to say that in actual experience these interpretants work in a mixed, at times chaotically associative and unpredictable manner. The following is quite a banal example, the purpose of which is merely to show how the conceptual frame works. More singular and complex phenomena will be engaged and clinical examples will be offered in the following chapters.

Let us take a clinical conversation in which I take part as a therapist. During the conversation I am aware, in differing forms and degrees, of the unfolding process of communication between the patient and myself. At some point in the exchange I become aware that the feel of the conversation is tense. This is perceived through markers like vocal tone, stress, pause, gesture, proprioception (inner awareness of the position of the body) and facial expression, though at first, I am not aware of their separate workings. They come to me as a gestalt, as explained by phenomenology. \(^{66}\) *Phenomenologically* I may notice that my interlocutor’s voice

---

\(^{66}\) Some of these attributes are perceived in the interlocutor and some I perceive as occurring in myself. It is sometimes impossible to determine who is the source of events and features, as the
sounds tense, unflexible, highly strung and rigid (note the metaphoric use of sound as material). I may find myself captivated by my own experiential sensations caused by the speaker’s voice texture that may cause me some kind of irritation, and I become aware of some tension in my limbs and my neck, a sensation that I am rigidly positioned in my seat and perhaps that my breathing too has been affected. From a semiotic point of view, it can be claimed that these sonic markers act as indexical signs, pointing, or directing my attention to a state of affairs that is not necessarily conspicuous to an external viewer.

Perhaps a specific sonic phenomenon will then come to my attention and I will begin to scrutinize it in its *semiotic value*, as a characteristic vocal gesture of my conversational partner – for example, a recurring mode of avoiding stress at the end of sentences leaving them, so to speak, hanging in the air. I may begin to inquire into this recurring phenomenon as a gesture (the sonic gesture of “remaining suspended in the air”) that has meaning. So, while I may be quite absorbed in semantics (what is being said, that may or may not have an obvious relation to the tense feeling), I may want to inquire into the tenseness of the patient’s speech. In addition, I may have a sense that I am being addressed to as an authority figure, perhaps a forbidding teacher or father. Thinking about the patient’s speech in its *addressivity* exemplifies a dialogic approach to the tense feeling.

---

intersubjective event is shared and co-created. But for purposes of simplification I shall describe now only my perception of the interlocutor (patient).

67 Of course, there is no taxonomy for such meanings, and, in the spirit of contextualism, meaning will be arrived at after careful, sometimes prolonged, consideration of sign versus context, process, etc. But this doesn’t preclude the possibility of recognizing a sonic sign as such – albeit one whose meaning is still unknown.
Obviously, these impressions are not explained in my mind in a clean-cut manner by phenomenology, semiotics or dialogics. A tone of voice can be explained by any or all three or in some form of combination. Tension in the voice can be experienced phenomenologically as an irritating texture calling to mind sound's tactile effects; as a grating tone of voice combined with a suspension of the ends of sentences; as an indexical sign of an emotional state in the speaker; lastly it will be understood in its addressivity – as directed to me (because my utterances have precipitated in the speaker a tense response, or because I represent other addressees – perhaps certain authority figures – and dialogues which invoke tension in the speaker, realized in his sounds).

To sum up, from a phenomenological point of view, my concern will be with describing what is going on experientially. What is my experience of sonic phenomena and what is the form and structure of my experience? What do I sense in the now of the speaking voice of my conversational partner and in my own voice (and in the dialogue as unit of inquiry)? I will then explore the nature of my experience of our voices, hums, rhythms, pauses – the lived-sensations, affects and textures: What images and metaphors come to mind? What of this experience is specifically sonic, and how is it surrounded and intersected with the visual, gestural and linguistic realms? After a while⁶⁸, in the course of conversation I will be able to identify to some degree recurring forms, structures, motives that are being articulated and gradually present themselves to my attention in the course of the sonic event. These recurring forms, by way of repetition (the creation of invariants)

⁶⁸ It shall be remembered that usually, psychoanalytic therapy is a lengthy process, providing the partners with an opportunity to slowly develop recognitions and understandings. Nevertheless, these may also occur on more immediate levels and come about as swift, intuitive recognitions.
enable me to envision norms and expectations of processes (like vocal moods) and structures (like vocal personae). I can inquire into the nature and mechanisms of these motifs and structures: a certain manner of speech, a pronunciation, a typical slurring of words, a manner of interrupting — these can be attended to as signs. Finally, I attempt to understand the meaning of sonic signs in terms of their communicational impulse and directionality: how do this person’s sounds address me, reach out to me?²⁹ Who, indeed, is speaking to me and who do I become (who am I as addressee) through this address? In reverse, to whom am I talking? Who am I addressing and what kind of sounds are elicited in me by my interlocutor? Do I sound, for example, motherly, fatherly, authoritarian, sure of myself or unsure of myself? Do I elicit a response, do I feel understood or rebuffed? And furthermore, what kind of a relational form is created sonically? Is the relation (sonically speaking) one of “chase and dodge” (Beebe and Lachmann, 2002), reminiscent of a fugal texture? Or is it one of complementarity and “being with”, characterized by a harmonious feel? Does it have a recurring thematic repetition, evoking a rondo form? The possibilities are, of course, endless, but each sonic relation will have its own implicit and unique feel, informing the speakers of the state of the ongoing relation, much like a musical improvisation.

A last word must be said about sound’s linguistic and nonlinguistic value. As we have learned from Bakhtin, the refraction of dialogic relations is infinite, and dialogues take place not only between conversational actors but also within the utterance

²⁹ Morson and Emerson (1990) explain similarly Bakhtin’s idea of the process of understanding: “...[t] is a four-tiered process: first the physical perception, then its recognition [or, a phenomenological approximation], then a grasping of its significance in context [semiotic articulation], and finally – and this is the crucial step – “active dialogic understanding. This fourth step ... is implicitly creative and presumes ever-new, and surprisingly new, contexts”. (ibid., p. 99, italics and brackets are mine).
itself. I have above pointed out the inner relations between different sound components in the creation of speech and the inherent, ongoing negotiation between centripetal and centrifugal forces within the utterance. These can be discerned in the transaction between the sonic performative level and the semantics that is being generated through speech, producing various effects such as irony, double-voiced utterances, etc. in myriad shades, as well as nuanced dialogues between particular idiolects and stylistic genres (dialects) of speech. Thus, one of the essentials in the investigation of utterances is the apprehension of the vocal nonlinguistic levels as juxtaposed with discursive value and verbal content.

II.V Developmental Research and its Relation to the Present Conceptual Context

The topics of embodied perception, nonlinguistic semiotics and dialogic processes are central in the development of self-consciousness, theory of mind, primary and secondary intersubjectivity and primary musicality, as conceptualized by developmental research. In the last four decades, studies conducted in the area of infant research (as well as studies of subjectivity and neuro- and cognitive science), have obtained results that are strikingly resonant with Merleau-Ponty and Bakhtin’s philosophic ideas. Thus, Beebe et al. (2003b), Beebe and Lachmann (2002), Bucci (2001), Gallagher and Zahavi, 2008; Malloch and Trevarthen (2009), Pally (2001), Schore (1997), Stern (1985, 1996, 2004), Trevarthen (1999-2000), Zahavi 2004, 2005 and others have tapped the embodied, prereflective (sometimes called pre-symbolic) abilities of humans for establishing reciprocal affective exchange through vocal, temporal and gestural acts. Both Trevarthen (ibid.) and Stern (ibid.), arguing
for an “innate intersubjectivity”, have been able to demonstrate that even very young babies possess an active and immediately responsive conscious appreciation of adult communicative intention. In all these studies voice, listening, rhythm and movement are seen, beyond being the psycho-biological source of musical ability, as talents that are “inherent in the unique way human beings move, and hence experience their world, their bodies and one another” (Trevathen, 1999-2000). In the same vein, Beatrice Beebe and her colleagues' numerous studies of vocal-rhythmic coordination in primary dyads and adult conversational partners exemplify the innate human givenness, need and ability to “be with” another and know an Other through these embodied capabilities (Beebe & Lachmann, 1988, 2002; Beebe et al., 1988, 2003a, 2003b). The above research has impacted certain areas of psychoanalysis, especially the intersubjective and relational movements that have all centralized the actual analytic relationship and the here-and-now of the analytic encounter and exchange as the major lever for psychic change. Their innovative ideas such as “now moments” (Stern, 1996, 2004) and “enactments”\(^{70}\) are heralded by Merleau-Ponty's ideas.

Modern infant research has unequivocally posited the dialogical origins of the mind, thus endorsing ideas that were put forward by existential-phenomenological philosophers (such as Buber, Levinas, Husserl and Heidegger and Merleau-Ponty) and writers such as Bakhtin. These authors have illuminated the relational aspects that lie at the basis of language and of the meaning of sound as sign, i.e. the structure of addressivity. Recent developmental and neuro- and cognitive research

\(^{70}\) These terms will be explained in the next chapter.
scholarship demonstrates that the processes of dialogue, mutual recognition, sharing, empathy, self and mutual regulation that run through sonic-temporal exchanges are central in building a sense of self, agency, social sense, symbolic development and language. How are these manifested through the voice will be the subject of the next chapter.
Chapter III
The Voice in Psychoanalysis

III.1 Introduction

The voice in Western tradition is positioned midway between metaphorical and concrete meanings, intersecting body and culture, language and music, interiority and exteriority, subjectivity and Otherness. Voice is understood in various disciplines through its physiological, linguistic, expressive, dramatic and musical functions. It is a renowned trope of authenticity, identity, subjectivity, authority and presence. This chapter will address the voice as a sonorous object in the analytic encounter, elucidating the operation of the concrete, nonlinguistic voice in the clinical setting and in the creation of a dialogic semiosis. The voice, in this sense, includes:

e. The voices of patient and therapist in the course of conversation. These consist of voice components such as pitch, volume, timbre and duration, as well as nonverbal sounds occurring in the conversation such as sighing, crying, laughing, coughing, and the like.

f. Elements and factors inherent in vocalization: the individual’s unique organization and articulation of sound, time (rhythm), breathing and silence.

g. Vocal-temporal aspects of the dialogue such as vocal correspondence, matching and synchronization and modes of turn taking.\(^71\)

\(^71\)The voice’s perceptual counterpart, listening, is addressed separately in the next chapter.
The first section in this chapter provides a critical discussion of the voice in psychoanalytic discourse, considering both cultural and philosophic influences on psychoanalytic conceptions of the voice. As we shall see, verbal language acts as the dominant interpretant, or conceptual frame of reference, in the understanding of all manifestations of the voice. Not only speech-voice, but also nonlinguistic vocal expressions are measured mostly in terms of verbal language – the quintessential example for which is, as we have seen, the conceptualization of nonverbal vocal expression as “para-linguistic”.

I will argue that psychoanalysis is pervaded by a conceptual rupture between the linguistic voice and the voice in its “other” (expressive, musical, intersubjective, embodied) nonverbal manifestations. The dichotomy between verbal and nonverbal reflects Western culture’s position vis-à-vis the voice, in itself embedded in a more fundamental psyche-soma split and a powerful philosophical visuo-linguistic bias (Ihde, 2007; Ong 1982/2002). The linguistic voice is traditionally considered as the stronger pole in this dichotomy, while the weaker, nonlinguistic voice is relegated to the margins of linguistic structure, unexplainable and ineffable. In this context the voice is considered as a regressive symptom, an archaic excess, or, when aestheticized as in singing, even perceived as a fetish (Bunker, 1934; Dolar, 2006). Admittedly, this predicament echoes a similar situation in the fields of music and aesthetics, where sound and music were repeatedly characterized through their resistance to verbalization. However, I have suggested that the long established view of nonverbal vocal elements as ineffable is more indicative of the shortcomings

72 These fields have been themselves under the influence of these cultural biases, and have in the last decades undergone profound changes in orientation with the advent of poststructuralism and postmodernism.
and limitations of the visuo-linguistic fundaments of Western culture than of the opacity of the sonic realm. As explained in chapter I, the logocentric position privileges spatial (visual) and conceptual aspects of perception and understanding, favoring distinctness and discreteness, designating objects (whether discrete or not) that are nameable, have clear contours, are extended in space and are capable of being represented symbolically. In this state of affairs, certain aspects of human experience have been marginalized, especially those that are inherent in present, embodied experience and resist verbalization. These aspects are often continuous, diffuse, gradient and interpenetrating: present-time experience and consciousness, music and the auditory-vocal sphere are prominent examples. The lack of a proper language to speak about these phenomena requires the use of adverbs, adjectives and metaphors to describe them (for example, in speaking about tonality of color and texture of sound). Recent cognitive and neuro-scientific research have maintained that this lack reflects the inherent gap between different levels of experience – like the verbal and nonverbal - and that, essentially, experience or music cannot be verbalized because of this intrinsic structural difference (See, for example, Bucci, 2001 and her concept of multiple code processing). Investigating and expanding our awareness to include these “wordless” phenomena and levels of experience could contribute to a more profound recognition- and thus conceptualization and integration – not only of these unnamable spheres but also of the totality of human experience. Indeed, psychoanalysis has greatly benefited from recent research in development, neuro- and cognitive science. Throughout recent decades, these fields have significantly expanded and enriched the

73 Yet in some cases it was privileged as being above and beyond language – perhaps the only realm that has acquired this status (see for example Schopenhauer, 1819/1966).
conceptualization of present time experience (Stern, 2004), subjective and intersubjective experience (Stern, 1985, 2004; Beebe and Lachmann, 2002; Malloch and Trevarthen, 2009; Trevarthen, 1999-2000; Zahavi, 2000; 2005), and emotional experience (Bucci, 1997, 2001; Pally, 2001; Rose, 2004; Schore, 1994). Similar endeavors exist in the field of music theory in an effort to re-conceptualize musical experiences which were previously considered ineffable (Cone, 1974; Cumming, 2000; HaCohen, 1999, 2001a, 2001b; Hatten, 1994; Keil and Feld, 1994; Lidov, 2005; Shepherd and Wicke, 1997; Tolbert, 2001, 2002 and others). These bodies of work corroborate earlier notions stemming from phenomenological philosophy and its way of explaining lived, embodied experience (Ihde, 2007; Merleau-Ponty, 1945/2008; Welton, 1999).

Consequently, the second part of this chapter suggests an alternative conception of the voice, one which reflects and integrates other views coming from phenomenological philosophy (especially Merleau-Ponty, ibid.), dialogic theory (Bakhtin, 1981), music theory and developmental research. The integration of these views, as elucidated in chapter II, sheds light on the role of the voice in the creation of semiotic fields and in the constitution of intersubjectivity. These views are fundamentally opposed to Western culture's long standing mind-body split in their endeavor to understand the human being as an essentially dynamic, nonlinear and semi-open system in which different modes of perception, cognition and communication interrelate in a continuous process of dialogue and integration. As an essentially dialogic and intersubjective process, the human being is dependent on and constituted by meaningful interactions with others.
Utilizing these ideas, I attempt to reestablish the position of the nonverbal voice in psychoanalysis and ultimately argue that the voice unfolds in a multidimensional manner, generating meaning in both linguistic and nonlinguistic modalities simultaneously. As I shall demonstrate, these two modes employ essentially different mechanisms of signification, constituting the process of listening and interpretation as an ongoing integration of this “multiple-code system” (Bucci, 1985, 1997).

In the third part of this chapter, drawing from music theory and, in particular, music semiotics, I discern various mechanisms involved in the creation of nonverbal vocal semioses, and suggest ways to segment and articulate the fluid, continuous nonverbal vocal sphere. Subsequently, I identify some of the forms and phenomena of the non-linguistic voice as it unfolds in the clinical psychoanalytic encounter, drawing on music theory’s study of musical fields. In this, music theory is indeed unique since it is the only discipline that deals with bare (i.e. nonlinguistic) sound as material for aesthetic-symbolic signification. The conceptualizations developed in this field, highly valuable in articulating the unfolding of the voice in the clinical psychoanalytic encounter, are accordingly adapted for this purpose.


Within the psychoanalytic discourse on the voice I identify three central positions, or “hearings”, of the voice. The ideas proposed by the authors included in each hearing are not uniform, and represent individual and sometimes different interpretations of
the voice. However, I have grouped them together according to their fundamental theoretical orientation regarding the voice: a. the voice as symptom b. the voice as object c. the voice as dialogic utterance.

First, one should note that although a certain correspondence between the different positions regarding the voice and the historical-conceptual course taken by psychoanalysis is to be expected, this correspondence is by no means total. Contemporary writers on the voice may embrace classical psychoanalytic tenets, while present-day ideas can be discerned, albeit in emergent state, in the thinking of early writers. In short, rather than providing a chronology of the voice in psychoanalysis, the following offers a critical evaluation of the voice, cutting across the historical span. Second, the three positions are not mutually exclusive and the different approaches often overlap. Finally, while some authors are more radical in their outlook, reflecting the profound changes that took place in analytic theory in the past century, most texts are theoretically layered and integrate earlier, more classical views with contemporary thought (see for example Bucci, 1997).

As a starting point for this discussion, I posit a distinction between two aspects of the voice which I designate as the **linguistic** and the **nonlinguistic**. The **linguistic voice** is employed in the production of verbal language, i.e. phonemes, words, and sentences. This is the voice used in the creation of socially agreed upon, symbolic language. It is important to see that the linguistic voice employs only certain vocal components, specifically those required for the production of phonemes. These are

---

74 These are coterminous with “verbal voice” and “nonverbal voice”. However, their designation as linguistic and nonlinguistic allows for the inclusion, and at the same time of the distinction, of the third category (paralinguistic), thus widening the scope of the possible roles of the voice, as opposed to the dichotomous verbal/nonverbal binary.
the elements of voice that can be reliably reproduced, identified, and repeated, thus complementing centripetal, homogenizing social forces.

The *nonlinguistic voice*, (or simply “Voice”), is that aspect of vocal production which does not directly contribute to the creation of discursive, lexical signification. It will be remembered that, in chapter I, I have distinguished between the paralinguistic voice and the nonlinguistic voice. I deemed the *para-linguistic* aspect as more properly pertaining to the realm of the linguistic voice. I have maintained that the category “para-linguistic” reflects a logocentric bias, but moreover, that it does not cover the entire range of nonlinguistic vocal phenomena occurring in conversation. For example, it does not account for aspects such as vocal texture and gesture, nor for global vocal aspects such as sonic atmosphere and persona (these will be explained in the next section). Finally, the para-linguistic category implies an isolated, “one-person psychology” conception that ignores the utterance (as conceived by Bakhtin) as an outcome of reciprocal, linguistic and nonlinguistic dialogic interaction and flow.

As I here argue and demonstrate further on, the nonlinguistic voice creates sonic semiotic fields which extend beyond the propositional, unfolding through mechanisms other than the verbal. Two central functions of the nonlinguistic voice are: a. the *signification* of the emotions and of energetic states, and b. the negotiation and signification of interpersonal forces. These two functions are performative and presentational, i.e. inherent in the lived present of the dialogue.

---

75 I employ the term signification, as opposed to expression. The latter implies a reified conception of an "inner self" that expresses itself outwardly through behavior. This conception has been critiqued both in psychoanalysis (e.g. Ogden, 1999a) and in music theory (e.g. Kivy, 1980). Similarly, I conceive of vocal phenomena as speech events that generate meaning dialogically through performance and continuous interpretation.
Although they have been acknowledged in psychoanalytic literature, they were not sufficiently elaborated or articulated.

I now elucidate three psychoanalytic hearings of the voice, followed by a critical evaluation.

**III.2.1 Voice as symptom**

A close reading of the literature maintaining this interpretation reveals that the speaking voice was acknowledged early on in psychoanalytic history – in fact, as early as “Studies on Hysteria” (Breuer and Freud, 1893-95/2000) - as an acute reflector of emotional states (Brody, 1943). However, as Brody warns, “this fact is so universally recognized that there is danger of it being taken for granted and overlooked. The analyst may tend to concentrate his attention solely on what the voice is saying” (ibid. p.371). And indeed, the scant literature that reflects this hearing of the voice indicates precisely the transparency that Brody warns against: The implicit understanding of the voice as a signifier of emotional states obfuscates the hearing of the voice as a sonorous - and psychoanalytic - object in its own right: when the voice functions “normally”, as the vehicle or material support for words, it is by and large not heard. In a Lacanian interpretation of the voice, the philosopher Mladen Dolar poetically remarks that if words are like the beads of a necklace (the chain of signifiers), then the voice is the transparent, invisible thread that binds them (2006, p. 23).

The voice in this interpretation resounds (becomes conspicuous) only when it departs from its verbal function - becoming a “disturbance” and giving rise to the
unequivocal opinion that “... disturbances of the voice and of voice production occur...as an outward manifestation of emotional conflict, in short, as a symptom of neurosis.” (Brody 1943, p.371). Similar hearings are offered by Bunker (1934), Ferenczi (1919), Freud (1893), Kohut (1957), Reik (1958), Suslick (1963), Tauber (1947) and others. The voice in these writings is fundamentally viewed as the locus of conversion and displacement processes, in other words as a bodily phenomenon that can be invested with unconscious ideation related to libidinal impulses. Consequently, as a symptom, the voice is conceived as a somatic representation of neurotic (hysterical or obsessional) intra-psychic conflicts. Other writers consider instinctual aspects of the voice and discuss the autoerotic component and the satisfaction of aggressive instincts through voice production (Greenson, 1954).

Paradoxically, the Voice makes itself conspicuous to psychoanalytic attention not only in its “disturbed” or “abnormal” manifestations, but notoriously, by its disappearance – as in cases of hysterical aphonia (Breuer and Freud, ibid., Tauber, 1947) and in the case of the silent patient (see early texts such as Bergler, 1938; Fliess, 1949; Levy, 1958). In fact, it appears that the phenomenon of silence coming from the patient, first regarded as a form of hostile resistance and later on considered as a plea for a new dyadic beginning and a benign regressive need (as in Balint, 1968; Bollas, 1987a; Coltart, 1993b; Winnicott, 1963/1984b), served inadvertently as a gateway for a new listening. This listening not only construed silence as a sonic substance which signification extends beyond the discursive, but also regarded the happening of the analytic situation as a lived event (see especially Khan and Coltart, ibid.).
The literature reports several extraordinary occasions in which the voice appears in nonlinguistic ways, causing the breakdown of linguistic structure, manifested for example in vocal or respiratory tics, hoarseness and stammering (see for example Tauber, 1947). Other unusual vocal events that are noted include singing or humming in sessions, the emission of a variety of nonlinguistic sounds (Hannett, 1964, Greenson, 1954), or unusual vocal presentations, such as a “feminine” voice in male patients (Ferenczi, 1919). These outstanding cases represent early discussions of the phenomenon of the voice in the clinical setting. The basic conception of the voice as a neurotic symptom appears to be widely consensual and is not further elaborated, perhaps with the exception of writers with an ego psychology orientation (see for example Kohut, 1957), that added the function of mastery and control as a characteristic of the vocal-auditory experience. It is telling that, as a rule, the voice in this hearing is the patient’s voice, while the analyst’s voice is hardly addressed, let alone considered as possibly displaying symptoms of its own.

From the point of view of the present research, the hearing of the voice-as-symptom is significant in that, though implicitly heard and virtually unexplained, the voice is still constituted as a signifier. It clearly points towards meaning, both in its “normal” functioning, as a medium for speech, and in its “abnormal” manifestations as a signifier of the repressed unconscious. I will, nevertheless, problematize this hearing and claim that its theoretic formulation contains an implicit pathologization of the nonlinguistic voice and is heavily biased by a series of implicit preconceptions.
III.2.2 Voice as Object

The second hearing, characteristic of writers with a relational orientation, apprehends the voice as related to the object (mother figure). Much like in the previous hearing, Voice is particularly noticed when departing from its verbal role, and is likewise heard as a sign of emotional disturbance. However, in line with these two-person psychology theories, rather than reflecting intra-psychic libidinal conflicts, Voice is mainly understood as embedded in the primary relation with the object, either representing the maternal imago, (i.e. the internalized experience of the mother), or conceptualized as an unconscious representation of the infant-mother relationship.

The maternal voice is here conceived as a powerful constitutive psycho-somatic experience, internalized at the beginning of life and consequently connected in the psyche to primitive unconscious fantasies of intimacy, closeness, merging and regression to early states of being or of return to the womb. These are evident especially in analyses of children and adults with severe psychological and characterological disturbances (for example Anzieu, 1979; Boyer, 1961).

As a consequence of this relational orientation, and in accordance with the conception of the analyst as a maternal object (in contrast with the previous conception of analyst as libidinal object), the voice of the analyst is more present and better theorized, and is equated with the maternal voice (Bollas, 1987a; Boyer,

---

76 As I have indicated in chapter I the term “relational” in this study (related to the concept of two-person psychology in the literature) includes all psychoanalytic theories that focus on relation as the primary psychic force rather than on libidinal intrapsychic impulse.

77 As Kaja Silverman argues (1988), the voice as representing the mother or the womb is a pervasive trope in Western culture.
ibid.; Winnicott, 1971). Nonlinguistic vocal and auditory phenomena coming from the analysand in the clinical encounter are largely interpreted as indicating regressive longings to preverbal states such as unconscious wishes of merging, being embraced, breast-fed, enveloped and cuddled (Bollas, ibid.; Greenson, 1954; Ogden, 1999a; Horton, 1984). Enjoyment caused by sound and music is often interpreted by these writers as containing regressive strivings to developmentally earlier stages of life (Bollas, 1992; Hannett, 1964). In contrast, vocal phenomena are also related to negative experiences of being penetrated, attacked, suffocated or castrated (for example Anzieu, 1979; Lecourt 1990). In the coming section, we shall examine the factors involved in these various experiences and interpretations of the voice, and what are the elements that makes the voice “given” as soft, cuddling, invasive, harsh and so on.

It is noted that this hearing of the voice has led to several important theoretical elaborations and is more richly developed than the first hearing of the voice as symptom. There have been several interesting conceptualizations of the maternal voice. It has been interpreted as a proto-object rooted in intra-uterine auditory experience and as a powerful vehicle in the expression of unconscious maternal phantasies towards the baby (Maiello, 1993, 1995). Touching upon the diffuse and unlocalized nature of sound, the maternal voice was also conceptualized as an early sonic envelope or an audio-phonic skin (Anzieu, 1979, Lecourt, 1990). Anzieu also defined the voice as a “sound mirror”, referring to early (certainly earlier than the well- known Lacanian mirror stage) maternal vocal responses and their effect on the baby’s emerging sense of self. Vocal manifestations in infants have been understood
as a transitional phenomena by Winnicott (1971), while Bollas offers an
interpretation of voice and music as transformational objects in both early and adult
life (1987a, 1992). These last two theorists focus on the voice as a representation of
the maternal relationship with the baby on one hand, and as an underlying psychic
source of creativity - sometimes connected with musical activity - on the other.\footnote{78}
Other object-relation oriented analysts view various vocal phenomena (like
repeating sounds made in speech or a characteristic tone of voice) as autistic objects
or as belonging to unmentalized or autistic-contiguous states, reflecting profound
primary disturbances in the sense of self (Mitrani, 1995; Ogden, 1989; Tustin, 1986,
1990).

This radically different – and diverse - conceptualization of the voice emerges as a
result of several developments in psychoanalysis. The application of psychoanalytic
treatment to nonverbal populations (children, borderline and psychotic patients)
inevitably directed the focus of attention to nonverbal aspects of communication.
The development of new, interpersonally oriented theories such as object relations,
self-psychology, the ideas of the British Independents, and American relational and
intersubjective theories alongside the increasing interest in the actual therapeutic
relationship as the locus of new experiences and a lever of change all led, although
indirectly, to a new hearing of the voice as a phenomenon occurring in and emerging
from a dialogical template. In addition, the cultural linguistic turn in the twentieth

\footnote{78 The present scope does not permit further elaborations on the reasons why voice can be
experienced as a transitional or transformational object. In brief, however, this means that the voice
may be chosen by an infant or an adult, to symbolize the early maternal relationship in the period of
separation. The designation of a transitional object is a unique creative act on the part of the
individual. Thus, for example, a special fondness for humming and singing in infants can be
understood to function as the creative, soothing transitional object that both marks and offers
consolation in relation to separateness.}
century represented in psychoanalysis by Lacan, redirected attention to issues of desire, subjectivity and intersubjectivity as constituted by language in the analytic encounter, and to the phenomenon of the voice and its signification in the actual intersubjective human encounter.\textsuperscript{79} \textsuperscript{80} This hearing of the voice as object has unquestionably expanded the narrower symptomatic view of the voice as sonorous object. First, Voice is convincingly interpreted through an early object-relation developmental framework, i.e. largely normalized as a nonverbal, primary communicative phenomenon. Second, it seems that Voice as sonorous object is more clearly heard, in the sense that its specific sonic nature is better recognized. Thus, the unique acoustic properties of sound, like its ability to spread in space or traverse physical objects, is reflected in such concepts as the sonorous envelope (Anzieu, 1979) or the sound-object (Maiello, 1993). Third, the hearing of Voice as a transitional phenomenon (Winnicott, ibid.) or transformational object (Bollas, ibid.) restores an essential connection between voice in development, voice in the clinical setting and music, a connection that is practically nonexistent in the first hearing, where voice and music are taken to be two distinctly separate phenomena. Finally, the two-person psychology perspective, much less mechanistic than the first view, construes the voice as stemming from a relational experience and introduces the analyst's presence and her voice as a decisive factors in the clinical encounter. This

\textsuperscript{79} Lacan was initially interested in the gaze and the voice as phenomena that are both material and immaterial, stemming from the body and yet exceeding it. He considered both as paramount embodiments of the objet petit a, but it was the gaze that was ultimately privileged by Lacan and conceptualized as the paradigmatic instance of the Imaginary order, while the conceptualization of the voice was abandoned (Dolar, 2006, p. 39).

\textsuperscript{80} Although interested primarily in meaning as emerging from linguistic discourse, Lacan paid attention not only to referential content emerging from language, but to the actual play unfolding through the linguistic sounds and signifiers. However, this interest always remained linked to verbal meanings.
orientation becomes fundamental in the third hearing of the voice, to which I now turn.

**III.2.3 Voice as Dialogical Utterance**

Voice as dialogical utterance constitutes the third hearing of the voice in psychoanalysis and coincides with the great theoretic diversification that psychoanalysis has undergone over the past four or five decades. The various understandings of the voice suggested within this group emerge from a multidisciplinary dialogue increasingly maintained between psychoanalysis and other fields such as developmental research, cognitive science and neuroscience (see for example Beebe et al., 1988, 1997; Bucci, 1985; Fosshage, 2004; Pally, 1996, 2001; Schore, 1997; Stern, 1975, 1985; 1987; Stolorow, 1997; Thelen and Smith, 1994; Trevarthen, 1998, 2001). Other central disciplines in this psychoanalytic dialogue are philosophy, modern hermeneutics and linguistics (Benjamin, 1988; Canestri, 2000; Chessick, 1996; Clark, 1997; Frie, 2002, 2003; Killingmo, 1990, 1999; Makari and Shapiro, 1993; Meissner, 2006; Stolorow, represent this growing body of literature). This multidisciplinary discourse has attempted to locate psychoanalytic theory in a wider cultural-philosophic perspective and has promoted a vigorous re-visiting of the processes involved in the establishment of the sense of self as emerging from a dyadic, dialogic, intersubjective matrix. Rather than focusing on the individual mind and its intra-psychic processes (critiqued in contemporary psychoanalysis as being a reified positivistic construct), this discourse emphasizes dyad, relation, dialogue and context as ongoing processes in constant flux.

---

81 The present nomination of the voice as dialogical utterance is obviously indebted to Bakhtin’s ideas, an extended explanation of which is found in chapter II.
It is noted that several authors can be identified with the present second and third perspectives of the voice, as these manifest a certain overlap. Therefore, a differentiating principle is here in order: the hearing of the voice-as-object is “purely” psychoanalytic in essence, that is, it is a theoretical construct that conceptualizes the nature of inner, unconscious mental life and its attending fantasy-life. In contrast, hearing the voice as dialogical utterance dialogizes psychoanalytic theory with extra-disciplinary conceptions, encroaching, as it were, on the purely psychoanalytical. This interdisciplinary dialogue presents the voice as intersecting – thus unifying – the material and the mental, the embodied and the psychological, the concrete and symbolic orders. The body is here seen as a basic experiential stratum that must be considered in the understanding of all mental phenomena. Although this discourse does not always elucidate how these experiences are constructed in the unconscious as it is classically conceived, it nevertheless builds valuable bridges between experiential, observed phenomena, philosophical and linguistic conceptions and psychoanalytical constructs.

An important part of this psychoanalytic inquiry centers on the nature of experiential, nonverbal, nonrepresentational states. Attempts are made to identify the mental-affective mechanisms and processes employed to transform nonverbal experience into more complex symbolic structures (see for example Bucci, 1985, 1997; 201; Stern, 1985, 2004). The implication of this investigation is a decentering of verbal language and the integration of other levels of human experience into the psychoanalytic realm, resulting in a redefinition of the psychoanalytic encounter and its various themes (such as the nature of self and subjective experience, the
transference relations, affectivity and the unconscious, the verbal and nonverbal domains etc.). It is important to emphasize that these conceptions are by no means uniform. However, they do share an underlying multidimensional and contextual perspective which not only opposes positivistic, binary and exclusively linguistic traditions, but also attempts to mend these pervasive cultural mind-body splits.

Since it is beyond the scope of the present research to fully delve into these varied theoretical elaborations, I will focus only on the central premises these share vis-à-vis the voice. The fundamental common ground here is that the voice in both its linguistic and nonlinguistic aspects, is heard as an ongoing, dialogical utterance formed and continuously unfolding through a meaningful interaction with an Other (the Other consisting of internalized figures, real persons and the analytic interlocutor). This interpretation emerges out of a broad and integrative conception of the human being as a semi-open system that is in an ongoing process of negotiation between multiple and essentially different forces and levels of existence such as psyche and soma, verbal and nonverbal, self and other, individual and social, past, present and future, etc. Rather than being heard as an isolated phenomenon identified in the individual speaker, a symptom one has, or as pointing towards a distinct mental representation, the voice as dialogic utterance emerges as a simultaneously presentational and re-presentational phenomenon that is multi-determined, vital and creative. The dialogical voice is understood as a present-moment event that unfolds in a dynamic context whose dialogical nature is infinitely refracted in myriad domains (somatic, mental, emotional, individual, dyadic, social), all present in every single utterance. Thus, vocal utterances constitute, enact,
negotiate and simultaneously amalgamate the subject’s linguistic discourse through individually unique performances, dialogic impulses and situated social and historical forces.

Some of the writers advocating this hearing address the voice explicitly (for example Bady, 1985; Killingmo, 1990; Knoblauch, 2000; Kristeva, 1986a, 1986b; Paul, 1989; Rizzuto, 1995, 2003; Stein, 2007) while others include it in a more general discussion of the nonverbal dimension in communication. In any case, all of these writers consider vocal-temporal components as central to primary communication and to the establishment of an early sense of self and self-and-other (see especially Beebe et al., 1988, 1997; Beebe and Lachmann, 2002; Maiello, 1993; Malloch, 1999; Malloch and Trevarthen, 2009; Pally, 1996, 2001; Rizzuto, 1995; Robb, 1999; Rose, 2004; Stein, 1999, 2007; Stern, 1985, 1996; Trehub, 1990; Trevarthen, 1999-2000, 2002). Voice is believed to be one of the fundamental means in human development for establishing primary affective relations and for the early constitution of subjectivity and intersubjectivity. Formed through the parameters of vocal pitch, texture, time and intensity, vocal dialogues between babies and caretakers are demonstrated, through empirical research, to be extraordinarily nuanced, displaying striking interactional feats such as self- and mutual regulation and responsiveness, synchrony, fine-tuned turn-taking and theme-and-variations vocal games as well as accurate vocal cross-modal matching (i.e. correspondence between vocal and visual, vocal and facial, vocal and gestural modalities).
In addition, Voice is demonstrated to be a powerful means of generating not only affect-laden intersubjective events and states of companionship, but also nonverbal semiotic fields that operate largely through specific organizing principles (such as repetition). The rich interplay between vocal components such as pitch, intensity (volume), pause and timbre, as well as rhythm and gesture lies at the basis of the creation of more complex vocal semantics through vocal gestures and narrative sonic envelopes. These are precisely the principles operating in music.

Several writers have attempted to link these ideas with adult psychoanalytic treatment. It is contended that the nonverbal, implicit principles of self- and self-other organization and regulation characterizing the primary dyad continue to exist and activate adult partners, mostly in an out-of-awareness mode. This implicit level of interaction is argued to be a central force in the therapeutic process (Beebe and Lachmann, 2002; Beebe et al. 2003a, 2003b, 2004; Bucci, 2001; Canestri, 2000; Knoblauch, 2000; Lachmann and Beebe, 1996; Morgan, 1997; Pally, 2001; Priel, 1997; Rizzuto, 1995, 2002; Rose, 2004; Stein, 2007). In an effort to link this idea to the therapeutic setting, some analysts examine interactional sequences in the present moment of the therapy, such as vocal-temporal and dynamic contours of the dialogue. Relying on non-linear, dynamic systems theory (Thelen and Smith, 1994), patient and therapist are viewed as two semi-open systems that are in a state of continuous motion and mutual responsiveness, regulation and adjustment, resembling, in many ways, musical dynamics, whereby the participants’ voices interact in a continual process of dialogue, reciprocation, reorganization and mutual

\[^{82}\text{This principle and others will be elucidated in the next section of this chapter.}\]
\[^{83}\text{The psychoanalytic signification of these levels of interaction is largely debated in the literature, as discussed in chapter I.}\]
reframing. Some conceptualizations emerging from these groups – especially coming from the work of Stern et al. (1992) - have had great impact, were effectively integrated in contemporary psychoanalytic discourse and are here especially pertinent because of their affinity to music, specifically Stern’s concepts of affect attunement and vitality affects (1985), the “moving along process”, “moments of meeting” (1998, 2004), the “pre-narrative envelope” (1992( as well as the concept of the “present moment” (2004). Other important conceptualizations coming from Beebe and her collaborators are those of self- and mutual regulation processes (Beebe et al. 2002, 2003a, 2003 b) and Pally’s concept of reciprocal responsiveness (1996). Analytic authors who engage in dialogue with philosophy and linguistics display a similar recognition of the role played by the nonverbal voice in speech. Different paradigms are offered to accommodate and integrate the various levels of communication including the signification established by the nonverbal aspect, emphasizing the performative level of speech. Canestri (2000), for example, informed by the linguistic theories of Apel and Buhler, proposes a pragmatic paradigm of linguistic activity in psychoanalysis that assigns a dominant place to the process of the production of the linguistic object and not only to its linguistic signification. Killingmo (1990) investigates the defense mechanism of isolation within the production of speech through the conceptualizations of the linguist Ivan Fonagy and the psychoanalyst Julia Kristeva. The latter, through her well-known

---

84 This mutual interaction takes place through embodied, gestural and facial axes as well. In classical psychoanalysis, though, where the body and the gaze are largely neutralized, the vocal-auditory channel is paramount.

85 These concepts will be explained in the third section of the present chapter.
concept of the semiotic chora (1986a), offers a hearing of speech that retrieves the nonlinguistic vocal interaction between infant and mother and locates it as an embodied semiotic activity that underlies linguistic discourse through life.

In summary, Voice as dialogical utterance is conceived from a fundamental position that highlights the nonverbal and enacted dimensions of experience. Whether coming from a neuro-scientific, cognitive, linguistic or philosophical perspective, the body and embodied experience are viewed as central interpretants in this hearing. Rather than viewing it as primitive, archaic or regressive, the embodied nonverbal dimension is conceived as a fundamental, non-pathological, non-regressive substratum of experience that, although occurring at an earlier stage in development, remains a fundamental component of experience and communication throughout life, supporting later, more symbolic, interactions and developments of psychic life. In addition, the nature of the nonverbal domain is increasingly understood as autonomous or, in other words, as operating through other principles and mechanisms than verbal communication, though continually interacting with it.

On the whole, human communication is seen as an ongoing integration and negotiation between various levels of perception, cognition and interpretation, which operate fundamentally differently from each other.

It is also important to note that Voice in this hearing is increasingly understood as a processual event. This corresponds to a general revision taking place in psychoanalytic thought in the last decades regarding traditional psychoanalytic constructs (for example the unconscious, or the sense of self) that are now seen as static reifications. An alternative understanding is emerging (see, for example,
Mitchell, 1993) that views these as dynamic, nonlinear processes that are in continuous flux and interaction with inner and outer factors, constantly going in and out of balance and always contextualized. This processual conception (that, yet again, is reminiscent of the unfolding of music) is essentially different from the static, spatialized approach that characterizes the previous hearings of the voice.

III.3 A Critique of Psychoanalytic Hearings of the Voice

III.3.1. Pathologization of the voice

The first problem regarding the psychoanalytic view of the voice stems from the linguistic bias of psychoanalysis. The centrality of language in the theory and practice of psychoanalysis needs not be further elaborated here, except for its implications regarding the voice. As Rizzuto (1995, 2002), Canestri (2000) and Frie (1999) among others, have pointed out, the classical linguistic paradigm in psychoanalysis has privileged the logos over the total significance of the speech event (Rizzuto, 2002) between patient and analyst. Despite the profound changes that psychoanalysis underwent throughout its history, this fundamental linguistic premise has remained unchanged.

In this light, my first critical statement is that Voice as sonorous object is virtually nonexistent in this classical paradigm: the heard voice is the word-voice, that is, a vehicle for speech production that, as Dolar (2006) succinctly phrased it, “goes up in smoke” as soon as discursive meaning is achieved. The role of the voice is merely to direct the listener's attention away from itself and towards verbal discourse -
narration, dream, memory. This is remarkable in light of the fact that psychoanalysis purposefully created the clinical arena as a speaking situation: it is indeed the talking-cure, not the word-cure, requiring the speaker's voiced presence. Furthermore, through the principle of free association and the analytic setting, the speaking subject in treatment is encouraged not only to let go of the discursive order and sequence of logos, but to enter an affective, dialogic “play area” (Winnicott, 1971), in which she is invited to find her metaphoric voice through the concrete use of the voice that expressively resounds in excitabile or monotonous, organized or disarrayed manners, laughing and crying, stammering and faltering, enacting the speaker's liveliness or deadness. The speaker is invited to play out the self, making it vocally present in an interplay with the analytic listener. And, as we well know, the analyst's voice and listening loom large in this psychoanalytic vocal play (more of this in the following).

And so, through its strict demands for abstinence (manifested in the prohibition of any form of physical touch, the restriction of bodily action, etc.) and the entry into the symbolic order, psychoanalysis paradoxically positions the embodied voice (representing the only possible form of physical action) as a central actor on the analytic stage. In this dramaturgy, vocal play can be heard as improvised music that, beyond its discursive import, immerses the analytic partners in its quasi-musical eloquence, melodic articulations, cadences and inflections that can be (only retroactively) understood as forming and making present the speakers' subjectivity and intersubjectivity. In this vocal mis-en-scene the voice performs, enacts and
creates emotional effects. The speakers-listeners are propelled into the vocal scene and are inevitably entangled in it, mostly unwittingly.

Evidently, this rich sonorous experience was neither explicitly recognized nor investigated in the symptomatic hearing of the voice. In fact, it is almost as if psychoanalysis was – neurotically? – defending itself against the Voice by means of its pathologization. Whether this exhibits a psychoanalytic resistance to the voice as a dangerous Other\textsuperscript{86}, or reflects a lack of conceptual tools to deal with the specifics of the phenomenon of sound as proposed above, or a combination of the two, remains to be elucidated.

So it was, that only through departing from its verbal role – by displaying “disturbances” or breaking up linguistic structure – Voice came to the focus of analytic attention. It was then pathologized, as we have seen, through two central perspectives: the one, predominant in the first of the three hearings, was treating it as a symptomatic disorder indicating a neurotic symptom\textsuperscript{87}. The other, characteristic of the second hearing, was relating it to a regressive – thus pathological – pull or longing to the maternal object.

Although correct up to a point, these interpretations are here argued to be limiting and reductive, and more importantly - relying on problematic implicit assumptions. While I am wholly in accord with the contention that the voice is a vehicle for the expression of emotions, and that emotional conflict, whether conscious or

\textsuperscript{86} Because of its diffuse, engulfing nature and its association with the maternal, feminine order, as proposed by Silverman (1988).

\textsuperscript{87} As discussed in chapter I, a similar interpretation was made regarding the artistic creative act. This interpretation was consequently heavily criticized as reductive (see for example Milner, 1987; Rose, 1980).
unconscious, is more often than not exquisitely articulated through the voice, I maintain that a conception of the voice as a symptom is only a partial interpretation that disregards its total signifying capacities. Especially problematic is the reliance on an implicit value system of normality and abnormality, where certain functions – for example verbal expression - are taken to be landmarks of normality and maturity, and any manifestations exceeding these are seen as pathological.

I raise a similar objection in relation to the second interpretation, that of the voice as object. Here, as above, I am basically in agreement with the understanding of the crucial, constitutive effect of the mother's voice and listening on her infant, and with the idea that the continuous somatic-affective exchanges (of which the voice is a central element) between mother and infant constitute a key factor in the formation of the infant's self and self-and-other relations. However, I view the pervasive interpretation of the voice as indicating regressive motivations to primitive fusional states as likewise reductive and based on questionable theoretical premises. As argued by Dowling, (2004), the much elaborated psychoanalytic theory of regression, though useful in its ability to link present and past experiences and modes of mental functioning, contains an implicit value system that equates evolution with progression, and dissolution with regression. Even more positive views of regression as a possible benign phenomenon (Balint, 1968; Winnicott, 1954/1984c) are implicitly permeated by this preconception. Dowling suggests that this implicit value system has given rise to powerful, though unfortunate, assumptions of social, cultural, developmental and individual value as embodied in notions of “higher”, “lower”, “primitive”, “mature”, “archaic” and “advanced”. Accordingly, the
psychoanalytic privileging of symbolic and verbal capabilities reflects an implicit belief that “later is better”. This preconception may be seen as embedded in an even more profound value system, namely the widely criticized phallocentric bias whereby certain cultural mores (such as maturity, rationality, language) are identified with maleness and privileged accordingly. The preverbal voice is relegated to the weak pole of this dichotomous conception, alongside femaleness, motherhood, the idea of regression, fusional experience (i.e. dedifferentiated experience), and, not accidentally, music and embodied experience.

In the male logocratic regime, the voice must function with discipline, articulation and control. By operating through certain selected vocal components which facilitate the formation of distinct, iterable sounds (i.e. phonemes), the voice helps create the formal, overt and undeniable dimension of verbal communication, enforcing the social system's needs for sameness and homogeneity (this idea is advocated in different ways by various authors; see e.g. Bakhtin, 1981; Corradi-Fiumara, 1990; Ihde, 2007; Levinas in Kenaan, 2008; Stern, 1985; not all of whom necessarily relate this trend to phallocentrism). “Other” aspects of the voice that are qualitative, gradient and lack clear boundaries – i.e. are difficult to verbalize, reiterate or consciously register – are related to the emotional (nonverbal, implicit and therefore irrational) domain, interestingly described by Stern as representing the “unofficial version”, the deniable aspect of communication. 88 Dowling (ibid.) points out the dire results of such assumptions, evident in attitudes concerning cultural, sexual and social “correctness” as well as goals of treatment. I maintain a

88 These are precisely the aspects of the voice that cannot be registered in writing. This of course, has serious ethical consequences regarding the concept of Truth.
similar criticism against the view of the nonverbal voice as exclusively reflecting regressive strivings to so-called primitive, archaic or undifferentiated epochs of life.

This regressive fallacy concerning the voice may be caused by several factors. The nonverbal voice is indeed difficult to discern and conceptualize, both because of its elusive nature and because of the gravitational pull of the linguistic aspect of speech. Obviously, the powerful effect of the nonverbal voice is more greatly noticeable during infancy, being a central channel of communication for the mother-infant dyad while symbolic language is not yet in place. With the advent of verbal language, attention gradually gravitates towards the verbal voice. But, as is well established by now, the nonlinguistic voice does not disappear, nor does it become any less influential in communication: it simply recedes from conscious awareness and resurges, in other words comes to the focus of attention, at certain times, at which it becomes more conspicuous (see next section).

This understanding, based on the idea that earlier phases in the development of the sense of self are cumulative and layered, and are continuously negotiated and activated throughout life, is proposed by Stern (1985) and concurred with by others. Stern problematizes traditional psychoanalytic developmental models, namely those of Freud, Erikson, Mahler, and Klein, whose working assumption is a sequential progress of various developmental stages that occupy critical time slots in development. In this conception, each developmental issue characterizes a distinct age-epoch, has an ascendancy and a developmental task to be resolved, after which it eclipses, is expunged, and a transition to the next stage is made. In contrast, Stern and others see these sequential developmental stages as normative negotiation with
life needs and tasks that become cumulative layers of the sense of self and continue to operate — simultaneously and no less centrally — after more complex developmental stages have been reached. He gives the example of making love as a fully interpersonal event that involves all senses of the self (ibid., p. 30). In this light, nonverbal expressions occurring in the clinical encounter are not necessarily a sign of pathological regression, but, as explained by Stern, manifestations of different modes of articulation and communication of various senses of self and intersubjective contact.

Peterfreund (1978) and Klein (1980), (in Stern, ibid., p. 19), argue that these developmental models share two fundamental conceptual fallacies, characteristic of psychoanalytical thought, in that they are both pathomorphic and restrospective. Stern argues that, in an attempt to explicate the development of psychopathology these theories work backwards in time, positioning pathomorphically chosen issues of adulthood in a central developmental role. In addition, in an important critique of psychoanalytic models of infancy, Peterfreund (1978) discussed the adultomorphization of infancy, i.e. the tendency to view and evaluate an infant not from the standpoint of his world but from an adult perspective, implicit in the characterization of normal infancy by such terms as “disorientated”, “delusional”, “omnipotent” and “narcissistic”. Peterfreund discusses and explains the confusion that arises from such adultomorphizations, characteristic of psychoanalytic thought, claiming that so much of our understanding of infancy has resulted from reconstructions through adult analyses rather than from direct observations of infants. It is my view that these implicit preconceptions (retrospective
pathomorphism and adultomorphization) lie at the base of the above outlooks on nonverbal vocal-auditory phenomena.

**III.3.2 Whose Voice Is It, Anyway?**

The next critical point is raised by the question: Whose voice is it, anyway? Or, what are the epistemological foundations in the various interpretations of the voice? This question is related to the much discussed shift in psychoanalytic theory from a one-person psychology to a two-person psychology model, reflecting a dissent from the solipsistic, “isolated mind” model that characterized Freud and the classical psychoanalytic tradition. Through these new understandings, the voice in the more traditional hearings clearly emerges as a product of the isolated individual model that views the patient as a distinct, separate entity. In these hearings it is the patient that is principally interrogated, and the answer to the opening question of this section is that, undoubtedly, the voice is the patient’s voice. Focusing on the intrapsychic realm, the exclusive hearing of the voice as a symptom the individual “has”, presents the voice as an objectified, reified, static and visual entity that is strikingly estranged from the sensitive fluid, continuous, dialogic and ever-changing nature of the voice (as I have argued above, this hearing does have value, but only as part of a multidimensional perspective).

Moreover, the analyst’s voice is conspicuously absent from this discussion, and, in the rare cases where it appears, it is through its symbolic conceptualization as phallus or authority figure, according to which all vocal phenomena are mapped (conversely, in the limited, though fascinating, literature on psychoanalysis written by patients, the analyst’s voice almost always features as a decisive factor in the
treatment. Unfortunately, this theme cannot be further developed here though it seems a promising topic for further research. See for example Chernin, 1996). The analyst’s voice is never implicated in its own potential symptomatic display, countertransferential or otherwise, and is implicitly portrayed as normal, that is, devoid of symptoms and disturbances. Modeled after the image of a neutral scientist, the analyst features as mostly silent external observer who decodes and interprets the symptoms he is presented (this stance of listening will be further elucidated in the next chapter).

This situation certainly improves in the second hearing as the analyst's voice enters the arena, representing the maternal voice and the early mother-infant relation. It could be roughly said that in the second hearing, the voice in question is the analyst’s voice (qua maternal object). As we have seen, this conceptualization represents an important shift in outlook. However, because some of these conceptions (especially the Kleinians and the British Independents) are strongly entrenched in Freudian theory, the vocal exchange is once again – at least to some extent – pathologized through the idea of regression.

---

The classical voice of the analyst, with its detached tone and scientific neutrality, may in itself be considered a symptom. This may explain why it often features in popular lore as an object of ridicule (reflecting the aggression and hostility it rightfully arouses in listeners). This is interestingly evident, although obviously in grossly distorted form, in the cinematography on psychotherapy and psychoanalysis. Indeed, it would be a fascinating theme to study: the changes wrought in the vocal persona of the analyst throughout analytic history in light of the changes in analytic theory (from authoritarian, libidinal object, to maternal object, and again to fraternal–egalitarian figure - and their manifestations in the sonic sphere).
Furthermore, the psychoanalytic identification of the analyst's voice with the maternal voice, though convincing, is too sweepingly and pre-conceptually applied onto the vocal sphere. First, the analyst's voice is rarely investigated in its concrete operation - its corporeality, symptomatic presentation, or how it is continuously affecting the patient as well as affected by the patient in the concrete event of the clinical encounter (there are certain exceptions, in which authors turn a sensitive ear to their own voices, such as Beebe, 2004; Knoblauch, 2000; Ogden, 1999a; Paul, 1989). Second, there is an implicit authoritarianism (and at the same time, an infantilization of the patient) in this interpretation. What is highlighted in the literature, is the unidirectionality of the powerful effect of the voice - from analyst to patient. But what about the parallel and simultaneous effects of the patient's voice on the analyst? These are seldom reported in the literature (again, Knoblauch, Ogden and Paul [ibid.], are exceptions). Still, the implicit “immaculate perception” (Stolorow, 2003) of the unaffected analyst somehow persists in this conception of the voice (though this outlook has been greatly challenged by later intersubjective and relational analysts, for example by Aron, 1996). There is little allusion in the literature to the multi-directional effects of the voice and vocalization, i.e. as occurrences that not only have reciprocal effects, but also, as suggested by Beebe and Lachmann (2002), have auto-effects in terms of self-regulation.

Through the recent work of the developmentalists discussed above, we come now to understand the voice as a dialogic, intersubjective, processual occurrence. So the question Whose voice is it?, related initially to the patient’s voice and then to the analyst's (qua maternal object), can now be answered differently. It is the dialogic
voice or voices of both patient and analyst, jointly constituting each other through the performance, interplay and inquiry of present and past vocal-dialogical occurrences embedded in their voice. Bakhtin’s philosophical concept of the dialogical utterance (explained in chapter II) addressed precisely what most developmentalists determine through their research. Ogden’s concept of the intersubjective third represents the same direction of thinking, and is quite useful here. He conceptualizes the clinical interaction as a dialectic tension between three elements: the patient and the analyst as separate entities, which, in the intersubjective analytical encounter, generate a third intersubjective "voice" that subsumes, but is larger than, its individual constituents, to which it cannot be reduced. Through these various conceptions, I elaborate the voice in the present study, hearing it as a dialogical utterance, that, while inherently containing an Other within, also constitutes the Other and simultaneously negotiates its own individual force and agency.

**III.3.3 Problems in Articulating the Nonlinguistic Voice**

The third analytic hearing greatly recovers the status of the voice in the analytic encounter. Nevertheless, certain problems emerge from this group of writers. One central problem that arises in the literature is the need to make more specific and clear articulations of the voice, which could account for the inherent structural attributes of sound (such as continuity, diffuseness, simultaneity, gradience). Unquestionably, the work of developmental and neuroscientific research has

---

90 I tend, however, to disagree with Ogden’s dialectic interpretation and favor Bakhtin’s dialogic perspective that proposes an unfinalizability, or an unresolved incommensurability between the different forces operating within the utterance.
positioned the voice in a more central place, developing important and feasible concepts that concern the vocal-auditory sphere. However, visual and other distinct (measurable) phenomena that are here seen as easier to conceptualize still seem to be privileged by research. And indeed, it is easier to grasp and elaborate bodily and facial gestures (as in the work of Beebe, 2004), or more distinct (and quantifiable) auditory phenomena such as pauses (Beebe et al., 1997). But phenomena such as vocal timbre, for example, are much more difficult to conceptualize, and so are continuous phenomena like the sonic ambience that unfolds in communication. For example, it is evident at what painstaking efforts the analytic experience in the moment is described by authors such as Ogden (1999a, 1999b), in his attempt to elucidate the simultaneous occurrences of reverie (thoughts, memories, bodily sensations) within the analytic dialogue. Because verbal language has a sequential, linear structure, manifested quintessentially in writing, it encounters great difficulties in explaining continuous, gradient and simultaneous phenomena. Bolas for example (2007), takes recourse in the use of the symphonic score as metaphor for the discussion of simultaneous levels of signification appearing in the sounds of the analytic conversation.  

Moreover, important conceptualizations such as “vitality affects”, “affect attunement” and the “pre-narrative envelope” developed by Stern (1985), as well as the increased attention and allusion to vocal tone by numerous writers remain at a somewhat generalized level. They are described through their outer contours as formal envelopes, or through metaphoric or impressionistic descriptions, and are in need of further elaboration. Evidently, the more profound and fine-tuned

91 The meaning of musical metaphors in psychoanalytic discourse will be explored in chapter IV.
conceptualizations of vocal occurrences inevitably require the development of a special – perhaps altogether new – terminology.  

The second problem lies in the actual designation of a nonverbal domain. The dichotomous division between the verbal and nonverbal domains seems to be trapped in the very body-mind split it endeavors to correct (see Harris' critique on this issue 1997). If, as the contemporary conception suggests, the basic unit of inquiry is the dyad, the fact that within the primary dyad one of the partners is verbal (obviously the mother) must be considered as crucially – and linguistically - influencing the evolution of the relationship. Thus, the early dyadic relationship cannot ever be seen as purely nonverbal, since the primary intersubjective sphere always includes a verbal participant.

Similarly, the conceptualization of mother-infant interactive articulations as pre-symbolic is problematic in my opinion, again creating a dichotomization between symbolic and non-symbolic domains. I prefer alternative conceptualizations that view human communication in a wider, semiotic perspective, as suggested by the work of Kristeva (186a, 1986b) and other analytic authors that rely on the semiotic models of Charles Peirce, Ivan Fonagy, as we have seen. These conceptualizations position human communication on a semiotic axis, where the totality of vocal expression is always understood as a signifying function. This semiotic vocal activity is conceived as embodied, though it is not purely bodily (natural), but always cultural as well. As Lacan felicitously phrased it, the first infant’s cry, the *cri pur* that emerges

---

92 Similar problems are discussed in the impressive and convincing attempts of music semiotics to make more nuanced articulation of sonic phenomena previously considered ineffable (see especially Cumming, 2001 and Lidov, 2005).
out of biological distress, is immediately, at first hearing, transformed by the attentive mother into a *cri pour*, i.e. introduced into the dyadic communication structure (in Dolar, 2006). Admittedly, various semioses display different modes of signification with differing degrees of complexity: as we have seen in chapter II signs can be iconic, indexical and/or symbolic. Signs also differ according to specific sensory modes: they can be visual, like objects and letters, motional or auditory. Accordingly they may be static or processual signs, involving temporal factors for their unfolding.

In view of the above, I now turn to the next section, where I suggest a new, multidimensional model of the voice in psychoanalysis.

### III.4 A Multidimensional Hearing of the Voice in Psychoanalysis

#### III.4.1 Introduction

My conception of the voice in psychoanalysis is largely in accord with the third hearing of the voice as described in the previous section. Fundamentally emerging from a dyadic relationship, the voice as dialogical utterance is a nexus of various interacting forces: body and culture, language and music, self and other. I view the verbal and nonverbal aspects of the voice as continuous, rather than dichotomous, dimensions in communication. I believe nonlinguistic vocal communication to be an ordinary dimension of communication that underlies verbal dialogue. Although it develops before the acquisition of language, it does not disappear with the advent of speech and its manifestations are not seen as regressive or pathological. On the
contrary, it is the nonverbal aspect of speech and its performative and addressive impulse that infuses the dialogue with a sense of vitality, and imbues the impartiality of language with the speakers' own voice.

In the present section, I attempt to further elucidate the nonlinguistic dimension of the voice as it unfolds in the conversational event, addressing those aspects that were not sufficiently developed in the literature (as I maintained in the critical survey). However, it should be remembered that, although my focus is here on the nonlinguistic voice, I eventually see it as part of the entire speech event that occurs through language.

Since the frame of reference is the dyad as an interactive dynamic system, the voice is, in a sense, the voice of the dyad: no vocal utterance can be understood apart from its belonging to the dyadic analytic process. Yet, the process unfolds through the unique vocal contributions of the individual participants who are themselves separate and distinct systems. In this dialogue, the uttering subject is not wholly accounted for as a product of context, as the postmodernist version would have it. In fact, one of the leading forces within the dyadic interaction is the individual's striving for distinctive identity, agency and freedom of choice. I hold that it is largely through the living sounds of the voice that these are generated.  

93 The issue of the obliteration of individual agency in postmodern accounts is a central one in contemporary psychoanalytic discourse. Because postmodernism has influenced psychoanalysis in viewing the individual as situated and embedded in social contexts, the place of agency and individuality in this contemporary relational paradigm constitutes a central axis of debate (see especially Frie's [2002] elaboration of this subject). In agreement with Frie and other existentially and phenomenologically oriented psychoanalysts, and consistent with Bakhtin's theory of dialogue, the present study constitutes an attempt to elaborate a conception of agency and individuality in a relational context that does not
The “eventness” of interaction, as well as the aspects of performance and participation are central and crucial to my understanding of the nonlinguistic voice.

My interest is primarily on process and present time experience and action rather than on formal analysis of material content – but we shall later see that the vocal process generates its own kind of content.

My conception of the nonlinguistic vocal realm draws heavily from two bodies of knowledge: 1. Early vocal interaction between mothers and infants. This theme is central in the work of infant researchers investigating the early emergence and establishment of a sense of self from a dyadic, dialogic, intersubjective matrix. This kind of inquiry places much emphasis on experiential, nonverbal and nonrepresentational states. 2. Musical improvisation, and specifically Jazz improvisation, as an analogue for nonlinguistic vocal interaction.

It is noted that some developmentalists have already made the connection between the two fields and have addressed their striking similarities (e.g. Gratier and Apter-Danon 2009). A few psychoanalysts have also discussed these (especially Knoblauch, 2000). It is important to note that in the field of music therapy, musical improvisation is one of the central means of engagement in communication and surrender to postmodern notions of the subject. One of the forms in which the agency of the individual comes about is, as I maintain, the use of the voice in the clinical situation. A similar view (though more linguistically oriented) is found in Ogden (1998). More on this issue is found in chapter II.

94 These authors have a particular affinity to music or are themselves musicians. It is not surprising that their sensitivity to sound and spontaneous musical expression and interaction makes itself present in their listening and participating in clinical interactions, and makes them particularly susceptible to issues of timing, rhythm, tonal nuances and global sonic textures. Of course, this doesn’t preclude the attentiveness of non-musicians to these aspects.
interaction, and much of its theory is built on this interrelation (see for example Bruscia, 1987; De Backer, 2004; Nordoff and Robbins, 1977; Pavlicevic, 1990, 1997).

The vocal realm I describe should be listened to as “music”: Within the fluid, continual stream of vocal interaction, forms crystallize, interpenetrate, affect each other and dissolve. In the process of grasping these phenomena, one must keep a manner of listening that retains both process and fluidity, that can oscillate between reflective and experiential listenings, between rapidly changing horizons and micro and macro levels of occurrence constantly moving and interacting, and like music, always in the process of coming together and breaking up, only to recreate new configurations. The emotive music of speech, in both its concrete and metaphoric levels, highlights the dialogic complexity (that is quintessential in music), between flow and form, sensuous experience and articulation, process and content. Finally this music must be considered as a subsystem of the larger dialogic utterance, in which it interacts with the flow of verbal semantics and signification.

### III.4.2 Central Characteristics of the Vocal Nonlinguistic System.

1. **Parameters.** The vocal nonlinguistic system consists of the continuous stream of vocal communication and involves all vocal components: pitch, timbre, duration, rhythm, tempo, loudness, breathing, range, stress and pause. Through the overarching principle of correspondence (to be explained below) that motivates the speakers as well as through nonverbal vocal mechanisms of communication, the vocal forms and phenomena that are created segment and articulate the continuous
vocal stream. One form of segmentation, the phonemic (linguistic), will not be dealt with it here. The other form of segmentation generates nonlinguistic forms.\textsuperscript{95} To elaborate it, I will discuss vocal timbre, vocal gesture, pause, vocal persona, vocal duet and vocal ambience. These categories do not exhaust all possible configurations of sound, but they are deemed central in the psychoanalytic clinical arena.

2. \textit{Autonomy}. The nonlinguistic vocal system is only partially autonomous from the verbal level.\textsuperscript{96} Listening to it as a separate entity demands a certain effort of abstraction or imagination: because of the powerful pull of language on our attention, we are always-already conditioned to focus on linguistic meanings.

An additional difficulty is related to the fact that the structure and unfolding of nonlinguistic vocal communication is quite different from its linguistic counterpart. It is a continuous (versus discrete) process of communication (Knoblauch, 2000), occurring as a simultaneous bidirectional flow between the participants (Beebe, 2003). Rather than decoding speaking and listening sequences, as is the case in linguistic communication, the focus here is rather on the continuous, simultaneous co-action, experience and participation of speakers in the moment.\textsuperscript{97}

\textsuperscript{95} Again, I wish to stress that this does not refer to the paralinguistic aspect of speech (see chapter I).

\textsuperscript{96} This is evident in theatre: different theatrical performances of the same play produce different – at times strikingly different - experiences. While the words of the text remain the same, the\textit{manner} of their performance varies, to the point where characters can appear wholly different. This implies that the nonlinguistic vocal realm is partially autonomous and can be organized and articulated apart from the linguistic level to convey distinct meanings. However, one should note that this may also derive from different interpretations of the emotive/affective import of the text.

\textsuperscript{97} This does not mean that participation is symmetric: participants may not manifest the same amount or form of activity. Particularly in psychoanalysis, where communication is largely determined by the different roles of the participants, this is the case. One can think of a jazz improvisation as a good analogy. The rhythm section (bass, drums, piano) play a rhythmic, repetitive stable role, marking mostly the rhythmic-harmonic framework, while the soloist (say, a trumpet) plays
Remarkably, both levels of communication occur simultaneously. The listening process involves an ongoing integration - ordinarily occurring out of awareness - between the two. The iconic and indexical signifying functions of nonlinguistic vocal signs operate simultaneously with the symbolic signification that comes about through language. This makes the semantics of the total vocal utterance a veritable polyphony that includes symbolic (linguistic) value and, at the same time, iconic and indexical (nonlinguistic) value, unfolding together through different vocal components and principles of operation. The different values of signification are not necessarily coextensive – they often do not overlap in terms of form and temporality. For example, various discursive contents can unfold in a specific, stable vocal atmosphere that unfolds simultaneously with the verbal modality although it is articulated through different sonic principles, thereby creating autonomous temporal and timbral effects that are deployed at a different pace and form and encompass the sequence of the discursive articulation in a global manner.

In this polyphonic conception, meanings emerge from specific blends of different signifying functions and are continuously negotiated. Rather than dichotomously dividing signs into verbal/nonverbal, the Peircean classification, more suitable, in my view, to the matter at hand, enables a more fluid conception of the vocal sign, where different forms of signification are in interplay and compete for centrality.

Dependent on context, at times the verbal meaning is more salient (what the a free (at times dense and ornate, at others, sparse and minimalistic) solo. The roles are not identical nor symmetrical, but the bidirectionality of intent and correspondence is essentially present. This principle is elaborated later on.

The difference and distinction between the various systems of vocal communication is not clear-cut in reality: verbal language is not only symbolic and nonlinguistic vocal signs sometimes verge on the symbolic. The differentiation here is methodological.
speaker is saying is more important); at other times, the iconic and indexical meaning of utterances becomes more central (how she is saying something\textsuperscript{99}). Ultimately, as meanings are contextually determined, the utterance cannot be totally stripped off of either its symbolic, iconic or indexical levels of signification at the event of uttering. In psychoanalysis, the considerable gaps and contradictions between the different aspects of the utterance (for example between the discursive contents and the way they are uttered), are one of the main foci of interest and attention.

3. \textit{Specific forms of attention and consciousness.} Related to the difficulty in perceiving the nonlinguistic vocal realm is the fact that, on this level, communication is implicit, non-symbolic and takes place mostly out of awareness. As we have seen, these features have been universally recognized and discussed by in various fields such as infant research, neuroscience, cognitive science and subjectivity studies. It is a level of communication carried in an embodied, automatic, prereflective manner. This growing body of research suggests that an expanded awareness of the nonverbal vocal dimensions of interaction can enhance the therapeutic process, both in terms of enhancing the here and now process of interaction, as well as by illuminating important themes which permeate the patient’s communications but are not manifested in the verbal level.

4. \textit{A musical system.} I maintain that the vocal nonlinguistic realm is systematically organized in ways that are reminiscent of musical fields. It is governed by similar

\textsuperscript{99} As I have indicated in chapter I, the \textit{how} of speaking includes the \textit{paralinguistic level} that is more directly related to the spoken words, alongside the \textit{nonlinguistic level} which is partially independent from verbal semantics.
principles of organization and, much like music, the vocal forms that emerge exhibit hierarchical relations: simple, basic phenomena are the building blocks of more complex vocal configurations. For example, the basic elements of timbre, stress, pitch and tempo build up a vocal gesture, and vocal gestures are the building blocks of vocal personae. However, in the ordinary experience of listening, which is based on a mechanism of oscillating, selective attention, complex syntheses do not necessarily achieve primacy over simpler phenomena. The selective focus of the listener on micro- or macro-level vocal articulations (for example vocal inflection and/or the global sweep of vocal ambience) is determined by various factors, such as personal inclination, goals, interest, or the prominence of the actual vocal material. Either way, higher and lower level phenomena unfold simultaneously and demand primacy of attention according to context. At times, vocal ambience will become the most salient phenomenon to perception; at others, a sudden, surprising change in the speaker's tone will come to claim the listener's full attention. The auditory figure-background configuration is thus always in motion.

5. Recursivity. In addition to being a nonlinear dynamic system, the nonlinguistic vocal realm (as all musical objects) is a recursive system. This means that on all levels of hierarchy, forms and phenomena are built through the same organizational principle and each level is contained in the other. For example, repetition and change is a pervasive organizational mechanism that is found on lower as well as higher and more global levels of the system. A complex structure, such as vocal ambience, entails the repetition of specific phenomena (vocal gestures, phrases, tempo, vocal tone) whose structural specificity is itself created by the principle of
repetition. The design paradigm of the totality of vocal phenomena can be construed as a kind of “nested doll” or “matryoshka principle”, an infinite interplay of mirrors denoting a recognizable relationship of similar-object-nested-within-similar-object. However, in contrast with these spatial metaphors, sonic objects-within-objects are displayed to perception continuously and simultaneously.

6. In chapter II, the nature of nonlinguistic vocal signs was generally explained. They were centrally qualified as nonpropositional, presentational, and processual. To briefly summarize these traits:

a. Nonlinguistic vocal signs are nonpropositional - they do not display or render propositional content. They are largely motivated signs that signify motional (energetic), emotional and relational meanings that can be perceived and interpreted through embodied and affective interpretants. Affect, energy and relations are the signifieds of these signs.

b. Nonlinguistic vocal signs are presentational. The uniqueness of sonic nonverbal phenomena is that meaning lies, for the most part, in the phenomena itself. Meaning does not lie “behind” the sonic nonlinguistic sign, it is not elsewhere (the Derridean différance, or deferral between sign and signified, though not totally eclipsed, is here minimal). It is the living sounds of speech and their complex characteristics as an inseparable part of their aural presentation that we are

100 A matryoshka doll, also known as a Russian nesting doll or a babushka doll, is comprised of a set of dolls of decreasing sizes, placed one inside the other. This structure usefully denotes this “object-within-object” design paradigm.

101 As I suggested, it is helpful to imagine listening to the sonic nonverbal dimension as if one is listening to a dialogue carried in a foreign language - where one can often sense (and understand) certain affective, energetic strands in the conversation without understanding their discursive content.
considering here: vocal texture and gesture are not arbitrary expressions of apriori ideas that are transmitted by speakers, nor are they superimposed onto discursive contents as some kind of sonic seasoning; rather, they are presentational features inherent in the deep structure of the utterance, enacted and generated in vivo mostly prereflectively, through performance.

c. Nonlinguistic vocal signs are processual: Because of their temporal nature, sonic signs are perceived and understood not simultaneously but gradually, as they unfold over time. The passing of time, however minimal, is needed in order for the perceiver to grasp and interpret elements such as pitch, timbre, volume, contour as well as their contextual relation to the overall sequence. On a larger scale (in more complex syntheses or sonic sequences), the listener’s perception and understanding is contingent upon the sequential concatenation of sonic events (Levinson, 1997). It is crucial to understand this point, because it is in this that nonlinguistic vocal signs are radically different than visual and linguistic signs, which are inherently more static and usually have more clearly defined boundaries.\footnote{Derrida’s différance notwithstanding, linguistic signs are basically more synchronic (their meaning is obtained all at once with their sound), while nonlinguistic vocal signs are diachronic – they require the passing of time to unfold and be understood. As we shall see in the next chapter, processual signs involve corresponding modes of attention and ascription of meaning.}

Two additional qualifications of the model must be now made:

\footnote{Although words and sounds have a similar temporal-sequential unfolding, the linguistic order is more arbitrary and static: phonemic sounds are invariant and so are syntactic structures. This makes speech more characteristically stable, while nonlinguistic vocal signs are more motivated, and can be deployed in different manners, for example polyphonically, and in circular (repetitive) ways.}
d. Nonlinguistic vocal signs are generative. We have seen that vocal phenomena have a signifying function – i.e., they are a sign of something: e.g., the emotional state of the speaker or the state of the relation. However, vocal phenomena are not merely conveyors of information, “symptoms” of speakers’ states of emotion or energy, nor do they constitute a string of isolated, discrete meanings. They also act in an articulatory function – they organize and segment the continuous vocal stream, generating identifiable shapes and sequences of events. Even more importantly, vocal phenomena create and articulate the intersubjective field. Because they are sensorial, presentational and performative, vocal signs penetrate, affect and influence the affective and behavioral patterns of partners or, in other words, the nature and form of intersubjective participation.

This combination of articulating and signifying, or doing and meaning, is what performing acts are about: meaning is created through performing (doing), and vice versa.103 In the process of participatory performance, acts of meaning are constantly negotiated because the participants’ central motivation is to cooperate and “be-with”. I maintain that central function of the nonlinguistic dimension is the enactment and signification of a togetherness: By vocal acting, the partners are doing something in order to be together. At the same time, their vocal behaviors give them continuous feedback, signifying the state of their unfolding togetherness.

Within this shared endeavor, there is equal importance to the signification of

---

103 This reading of “doing” differs from the psychoanalytic understanding of various forms of acting as a “repetition compulsion” that comes about to unconsciously express ideas that have not been yet articulated in thought (Freud, 1920/1955). Interestingly, this conception of doing, or “acting out” as it is called, was somewhat modified in modern approaches that view these performances as “enactments” that constitute an important lever in the analytic communication (see Psychoanalytic Dialogues, (2003), issue 13(5) devoted in its entirety to the theme of enactment). Nevertheless, as we shall see, doing in the sonic dimension fundamentally involves repetition.
difference and individuality and this is achieved, as we shall see, through the nonlinguistic vocal articulation of change, discrepancy and slippage and the continuous disruption of togetherness.

e. Vocal signs are a nexus of past and present. The overall meaning of vocal phenomena (including phenomena occurring on the level of the individual) should be understood as fundamentally dyadic, dialogic and intersubjective, in two simultaneous senses: a. In the present tense of the clinical encounter - as a result of the actual analytic interaction between therapist and patient. b. In a historic sense - as manifesting past, internalized dyadic experiences and patterns of communication which are being reenacted in the clinical situation.

Thus, any given utterance expresses a continual interaction between past and present dyadic experiences. The intertwining of past and present is, of course, cardinal in psychoanalysis: it is the elucidation of the transference – the mapping of the past onto present experience - which essentially constitutes the curative force of the analytic encounter. In this context, I suggest that the transferential relation is also reflected in the nonlinguistic vocal interplay between the analytic participants (clinical accounts supporting this idea are provided by Killingmo, 1990; Paul, 1989).

Through micro-analytic methods of analysis, infant researchers have greatly contributed to the understanding of the active, ongoing interrelation of the two temporal levels (see especially Stern, 2004). It is one of my objectives to demonstrate how this polyphonic interplay between past and present is enacted in the vocal dimension.

* * *
Having qualified the proposed model of the nonlinguistic Voice, in the following section I present the specifics of the model.

I first elucidate vocal principles and mechanisms of organization. In this category, the focus is on communicational impulse and transaction, i.e. operational modes of vocalizing that are governed by dialogical, intersubjective motives. In the second category, I identify material vocal phenomena and forms that emerge through vocal activity. I begin by indicating individual forms (phenomena discerned in the individual's utterance: vocal timbre, pause, vocal gesture and vocal persona), and proceed with dyadic phenomena (vocal duet and vocal ambience).

**III.4.3 Principles of organization: The Rules of the Game**

*The Principle of Correspondence as a Motivating Force*

The overarching motivational principle governing the dyadic nonlinguistic vocal realm is the principle of correspondence. Different writers discuss, in different terms, the inherent motivation of human dyads to self-organize and develop towards increasing integrity, coherence and self direction. This is true about primary dyads (infants and mothers) as well as adult ones. Partners are inherently motivated to coordinate their experience through cooperation, negotiation of intents and regulation of the constant flow of internal and external input coming into the
Correspondence helps the partners to organize their communication by contributing to the capacity to predict and regulate their mutual ongoing responses. The mechanisms and phenomena discussed by different authors, like self- and self-and-other regulation, imitation, matching, attunement, sympathy, empathy, companionship, resonance and responsiveness are all based on the inherent motivation of conversational partners for the creation of correspondence. An extensive overview of the themes of correspondence and matching as elaborated by prominent developmentalists (such as Meltzoff, Sander, Stern and Trevarthen) and their relevance to adult treatment is found in Beebe et al. (2003a).

Stern and Trevarthen are two prominent researchers who consider correspondence as a central lever in the creation of intersubjectivity. They view correspondence as based on an inherent human need for a sense of togetherness, social belonging, knowing and being known and the desire to continually expand intersubjective communication – i.e. the capacity for psychic intimacy, companionship, openness to disclosure, permeability and interpenetrability of affective states of mind. According to Stern (1985), the human desire to know and be known in the sense of a mutual revealing and sharing of subjective experience is great. In fact, he maintains that this powerful motive can be felt as a need-state.

The term “correspondence” contains the mutuality of the dyad (co-) as well as the principle motivation of the dyadic partners to respond to each other. Webster’s Dictionary defines correspondence as a term that encompasses a. an agreement, suit, or match with something or someone (in our terms, a co-being, co-existing or a

---

104 The evolutionary and biological determinants of such motivations constitute an independent area of study that is beyond the scope of the present research.
togetherness) and b. the idea of similarity, analogy or equation with something or someone. This is endemic to the nonlinguistic sonic order. Whether we are considering musical objects or vocal dyadic communication, we are witnessing a striving for co-being, resonance and togetherness mostly occurring, as we shall see, through a mechanism of repetition. Correspondence, on one hand, encompasses all of the above mentioned modes of being-with, on the other hand, it is not necessarily confined to positive interaction. Confrontation, disruption, conflict, contrast and surprise, emotional distancing and evasion are also important forms of co-response, and constitute change, the counterpart of repetition.

The question is how is the principle of correspondence manifested in vocal action? What are the mechanisms that make vocal acts and phenomena cohere and give rise to a sense of communication and meaning? The answer can be found in the two mechanisms which help create a sense of correspondence: a. The mechanism of repetition and change. b. The mechanism of vocal narrative.

**The Mechanism of Repetition and Change**

Because the sonic nonverbal order is nonreferential, meaning is not generated through cognition, symbolization or intellectualization, but through the performative, concrete mechanism of repetition and change.

**Repetition in Music**

When something repeats, a cumulative effect is created: through repetition, phenomena are grouped or marked in perception as circumscribed entities possessing a recognizable shape. Throughout the dense, chaotic stream of
experience, such entities create invariants, and function as “islands of consistency”, to use Stern’s phrasing (1985), facilitating the organization of experience in various senses. In the sonic realm, invariant forms create expectancies, and expectancies mark the framework or boundaries of the sonic field and its range of behaviors and phenomena (the music theorist Leonard Meyer [1956] based his seminal theory of musical meaning precisely on this psychological phenomenon). As described further on, this is the very mechanism which helps dyads organize themselves in coherent, predictable ways.

Repetition is manifested in sonic fields in various ways, with varying functions and meanings. The music semiotician David Lidov (2005), indicates the relative concreteness of repetition: beyond questions of style and musical abstraction, some unit, motif or element recurs. The repetition can be long (the unit is repeated many times), or short. It can occur in lower and higher levels of the sonic field, and through different musical components – a small unit like a rhythmic beat can be repeated just the same as phrases or larger harmonic structures.

The functions of repetition can be syntactic, contributing to the creating of hierarchy, organization and segmentation, the identification of invariants as well as the clarification of style. However, as Lidov points out, repetition can have an affective value in and of itself, as a vehicle of drama, humor and poetic reflection, as a mechanism for creating hypnotic continuity (which is contrary to its effect of segmentation), feelings of pleasure and unpleasure and for building up tension. When repetition is sufficiently prolonged, it creates an effect of saturation, evoking feelings of boredom, stasis or circularity. Repetition also has a rhythmic
consequence, since it marks equal duration in time, thus enhancing rhythmic consciousness. This is especially evident in performing dyads – whether in mother-infant communication or in musical improvisation. Lidov distinguishes between types of repetition that can be informative for our purposes. These three categories of repetition are neither mutually exclusive nor exhaustive.

The first type is *formative repetition*, which refers or directs attention to and marks the material repeated. It fulfills syntactic functions such as defining the units of the sonic field and establishing their position in a hierarchy of more complex forms. A repeating vocal gesture, for example, becomes a building block for the more complex structure of vocal persona.

The second type Lidov identifies is *focal repetition*. This is a self-referential type of repetition in which the focus shifts from the material that is repeated to repetition as activity per se. Lidov considers this activity to be symbolic, since it elicits associations and invites interpretations: “focal repetitions in music have a considerable power to evoke feelings of situations typified by repetition: activities that go on and on, rituals, compulsive actions, getting ‘stuck in a rut’, emphatically accented speech...”(Lidov, 2005, p.34). This type of repetition is obviously of special relevance in the psychoanalytic arena, and evokes the Freudian concept of repetition compulsion (1920/1955). But whereas in psychoanalysis repetition is mainly understood as an unconscious, compulsive, even pathological form of action, in music, and as I claim, in the nonlinguistic sonic level, repetition is a central communicative device.
The third type of repetition, according to Lidov, is textural repetition. The effect of continually repeating an idea or motif releases the attention from the repeating phenomena, shifting it to other occurrences in the field. Thus, the ongoing repeating phenomena create the effect of background. As we shall see, this type of repetition lies at the base of global vocal phenomena like vocal duet or vocal ambience (see below).

It must be remembered, however, that Lidov discusses repetition as a compositional means within a highly specific aesthetic language - the language of music. Moreover, he is engaged in an a-posteriori analysis of pre-composed musical forms. Even more instructive for our purposes are the various functions of repetition in spontaneous communication within mother-infant interaction and dyadic musical improvisation.105

Repetition in Infancy

Infant researchers have universally demonstrated the centrality of repetition in early infancy as an individual and dyadic means for organizing the total flow of experience. Stern (1985) discusses the different uses of repetition in the infant’s self- and self-and-other experience. His seminal concepts of “vitality affects” and “affect attunement” (explained below), are largely based on the repetition which occurs in both individual and dyadic experience. Stern suggests that the infants’ initial global, unified experience is lived in terms of intensities, shapes, temporal patterns,

---

105 These effects occur in all musical improvisation (individual and ensemble). For present purposes, however, I will refer to the improvisation unfolding between two musicians, hence “dyadic musical improvisation”.

159
categorical affects\textsuperscript{106} and hedonic tones. In time, and largely through the fact of repetition, these experiences, recast as generalized, become invariant patterned constellations of experiences of self and self-and-other (for example, self experiences like hunger and tiredness, or self-and-other experiences of joyful playing or feeding). Infants gradually and systematically organize the elements of these experiences by identifying their invariant constellations – thus creating a sense of coherence and continuity in the experience of self and self-and-other. Stern proposes that whenever a constellation is formed, the infant experiences the emergence of organization. What is repeated or perceived as repeated are the internal affective and dynamic components of the situation – temporal shapes, contours of intensity, instances of emphasis or pause, etc. For example, the internal experience of hunger, repeated many times a day, has a dynamic shape of increased intensity - in musical terms we might call it a crescendo. Due to the factor of repetition, the infant learns to recognize it as a familiar feeling that has an identifiable contour of time and intensity. As Stern and others demonstrate through the important concept of amodal perception, the shape of a crescendo can also be experienced in other sensory or perceptual areas: one can experience the increase of tiredness, interest, curiosity, laughter, crying, dyadic excitement, etc. This shape of crescendo is similar not only across different experiences, but also across different sense modalities (we experience crescendos through visual, auditory, tactile,\textsuperscript{106}The term “categorical affects” refer to the Darwinian classification of affects into discrete categories: happiness, sadness, fear, anger, disgust, surprise, interest and shame and their combinations (in Stern, 1985). This discrete classification is different than (though it indeed contains) vitality affects (explained below).
olfactory and proprioceptive modes). Stern refers to these global, supra-modal shapes as vitality affects.  

Stern maintains that the global, subjective world of ongoing emerging organization – the emergence of invariants through repetition - remains the fundamental domain of human subjectivity throughout life. In phenomenological terms (see for example Zahavi, 2005), this is the domain of the first-person prereflective lived-experience, which is mostly nonlinguistic. Later on in life, this domain provides the experiential basis from which thoughts, perceived forms, identifiable acts and verbalized feelings will arise. Finally, “it is the ultimate reservoir that can be dipped into for all creative experience” (Stern, ibid., p.67), as all learning and all creative acts begin in this domain.

Repetition in Dyadic Interaction

Since the focus here is on dyadic interaction, it is of special interest to see how repetition works in interpersonal vocal communication. In infancy, the repetition of vocal forms is maintained by various researchers to be a central means in the creation of a meaningful communication, leading to a mutual sense understanding, belonging, and a shared sense of “happening” (Gratier and Apter-Danon, 2009). Stern (1985), Beebe (2005), and others similarly demonstrate that different forms of repetition in the caretaker’s behavior not only help the infant organize experience

---

According to stern (1985), these are called “affects” rather than “feelings” because they are not categorically and linguistically defined (as in “anger”, “joy” [see fn.14]). They are more diffuse and qualitative, denoting the quality of feeling: bursting, fading away, etc. The specific form of a categorical affect is determined by the vitality affects that compose the specific feeling. This idea is consistent with Suzanne Langer’s elucidation of the relation of musical affective meaning, as well as with 18th century music theorist Daniel Webb’s (1760/2003) understanding of music’s articulation of affect.
into invariants, but also underlie the creation of intersubjectivity, i.e. the creation of a sense of a mutual sharing of affective states of mind.

Repetition in a dyadic context involves acts of matching – one partner matches the other's state of mind or affect by repeating something. Repetition in dyads is complex and is the base of several dyadic forms of correspondence, like empathy\(^\text{108}\), and affect attunement (Stern’s seminal concept, denoting the nonverbal sharing of intersubjective states [1985]). \(^\text{109}\) As Stern has demonstrated, repetition in affect attunement is not a straightforward imitation of the other. In order to be emotionally effective, it must occur through the cross-modal repetition of abstracted elements that represent the affective quality of the shared emotional state. For example, attuning effectively to the other’s affect of exuberance can be done through the voice, through bodily gestures, facial display, and so on. It is of key importance to our purposes that what is abstracted and repeated are musical elements: intensity, temporality (like tempo or rhythmic pattern) and tonal contour.

Although these forms of interaction are construed by the different writers as presymbolic, it is nevertheless evident that the abstraction of qualities by amodal perception forms a beginning of symbolic activity. As Stern proposes, both vitality 

\(^\text{108}\) Some form of dyadic repetition – thus, imitation - underlies the phenomena of empathy and sympathy. Imitation and its relation to sympathy has roots in mythology, as seen in the myth of Narcissus and Echo (HaCohen, 2001b). The echoing voice is here shown to be a central means for conveying sympathy and being with, and was used as a major device in certain types of religious music as symbol of sympathy with a suffering or joyful persona. Yet, as this principle developed, it went beyond mere repetition to create more complicated forms of correspondence, and also to facilitate other ways of being with, more iconic and indexical in nature.

\(^\text{109}\) Stern (2004) defines affect attunement as “a special form of behavior in response to the communicative affective behavior of another. Just as imitation is a faithful rendering of the other’s overt actions, affect attunement is a faithful rendering of what the other must have felt like when he or she expressed him- or herself with those actions.” (ibid., p. 241). The referent of attunement is thus not the action of the other, but the feeling behind it.
affects and affect attunement involve a recasting of a subjective state in different terms and through different modalities, such as gestural, facial, bodily, vocal, and so on. Each manifestation of an affect has some degree of substitutability as a recognizable signifier of the same inner state or perception, hence, a primordial form of nonverbal metaphor and analogy (see above Lidov’s similar view of repetition in music). This, as Stern claims, is an essential step toward the use of symbols. This idea also reinforces the present study’s view that signs and symbols are not categorically differentiated, but share a mobile, fluid interface.

Gratier and Apter-Danon (ibid.) centrally address the elements of repetition and variation and compare the vocal interaction between infants and mothers to jazz improvisation, where repetition is central. By repeating rhythmic elements, motifs and other musical structures, improvisational partners seek and maintain a sense of coherence, synchronization, boundary, predictability and expectation. In turn, these sonic invariants possess an affective value, both subjective and intersubjective, arousing a sense of stability, safety, togetherness, belonging, companionship and mutual reliance (see Berliner, 1994 and Monson, 1996 who present numerous accounts of musicians’ positive emotional statements about their experiences of musical correspondence and cooperation). These affective outcomes obviously have paramount importance in processes of communication, whether in development, musical performance or the clinical encounter.

**Change**

Individuals and dyads, as open systems, largely organize their experience through the mechanism of repetition, through which they create invariant categories.
Nevertheless, the ongoing current of life experience continually introduces new occurrences in the system: physical, emotional, developmental, situational, inner or interpersonal. At the same time, changes are not only imposed by circumstances. As has been demonstrated empirically in infant research (for example Stern, 1985), when something is repeated for too long and in too precise a manner, levels of attention and interest decrease. Exact repetition results in a sense of stasis, monotony, boredom and, as evident in severe mental psychopathology, a sense of psychic deadness. Yet, as we have seen above, exact repetition in music does serve certain functions.

Changes in the system are effected not only because of the intrusion of new elements or because of an inherent need for variation. They occur as a default condition because individuals are differently organized systems. Conversational partners’ differences and gaps in style, motivation and expectations continuously disrupt the balance of the dyadic system, forcing partners to keep up an ongoing negotiation of their interaction. These disruptions are precisely what keeps the dyadic system vital, creative, in continuous search for new solutions, preventing it from becoming boring and mechanical. Dyadic mismatches (leaving a margin of difference, or introducing change) are considered, in developmental and music literature, an important and fertile area for growth and development.

In general terms, it could be said that while repetition introduces stability and predictability, safety and reliance, change introduces the negotiation of uncertainty.

As is well known by clinicians, monotonous repetitiveness is one of the behavioral characteristics of psychosis. Interestingly, it is this very monotony that serves in certain pathologies as a form of psychic defense, for example in autism or in early psychic trauma (Ogden, 1989; Tustin, 1986).

The important Kohutian concept of “empathic failure” (Kohut, 1959; Kohut and Wolf, 1978) can be understood as a form of mismatch, that, when acknowledged, has important curative effects.
and novelty, mobilizing and developing coping mechanisms, agency and creativity. If they become too extreme, repetition can be negatively experienced as rigid, monotonous and tension provoking; and change, as chaotic, bewildering and similarly arousing anxiety. It has been demonstrated (for example by Beebe and Lachmann, 2003a) that a medium range of these is optimal. A successful dyadic negotiation of change results in a new form of correspondence, and an expanded sense of intersubjectivity. Through the concept of “moments of meeting”\textsuperscript{112}, Stern et al. (1998) maintains that new forms of relating discovered in the course of communication creatively expands the repertoire of behaviors and experiences, enhancing the sense of mutual reliance and intimacy.

In music, where repetition is so pervasive, change is its necessary counterpart. Ranging from minute nuances to radical contrast, change creates a large range of effects. Because of sound’s temporal nature, changes can be gradual (like the effects of crescendo and diminuendo), creating dramatic effects of mounting or decreasing tension. Changes can also be drastic – creating effects of contrast, suddenness and surprise. In addition, the polyphonic nature of the sound can simultaneously perform repetition and change on different levels: while one of the sonic elements changes, others remain constant (a well known example is Ravel’s \textit{Bolero}: While the melodic phrase and the rhythmic structure repeat continually, creating a hypnotic effect, other elements - timbre, density, orchestration and

\textsuperscript{112} Stern (2004, see also Stern et al. 1998) defines a “moment of meeting” as “a present moment between the analytic participants that potentially resolves a crisis created by (a previous) present moment.” It is the innovative, sensitive, authentic response of the therapist in the moment of interaction that can lead to a successful moment of meeting, thereby reshaping the intersubjective field and expanding its relational possibilities – hence the important role of change in this encounter. In this light, though in a different context than the clinical, the powerful emotional effect of musical improvisation can be understood as relating to successful moments of meeting between the musicians.
volume – change, creating an effect of mounting tension). The simultaneous interplay of repetition and change is most characteristic in Jazz music, where improvisation (i.e. change), unfolds in relation to a basic melody and an invariant harmonic structure. Similar effects are found in the clinical situation. For example, a speaker’s vocal timbre may change, but her tempo (velocity of speech) will remain constant. In the analytic dyad, one of the speakers (usually the therapist), will have the role of creating stability through his use of the voice, while the patient’s role facilitates greater freedom of expression (the analogy of the analytic dyad to a musical soloist and accompanist will be referred to later, in the section on vocal duet).

Finally, repetition and change are the underlying principle of the creation of individuals' and dyads' vocal genres and specific styles, or, as they are often called in music theory (Hatten,1994; Lidov 2005), musical dialects and idiolects. Vocal styles, as repeating, stable patterns of utterance, are related to specific idioms and a specific “aesthetics of being” (Bollas, 1987). It is through these genres that issues of identity, individuality, choice and preference are further manifested. Vocal dialects and idiolects also indicate the relationship and negotiation between individuals and their general cultural context as well as the style in which they choose to manage the present moment situation. The negotiation between dialect (social norm) and idiolect (personal choice and interpretation) is always idiosyncratic and offers a window of opportunity for individual choice. Listening to the vocal styles of individuals and dyads in the clinical situation in terms of dialect and idiolect enables a more nuanced, sensitive perception of important issues – especially those
regarding the relationship between the individual and the outside world - as they are instantiated through the voice.

**Sonic Narratives: Shared Vocal Voyages**

Another sonic process generated by the motivating principle of correspondence is the phenomenon of nonlinguistic vocal narrative, or as I refer to it here, shared vocal voyages. The continuous vocal co-action of partners leads to the formation of units and phenomena that interconnect and concatenate in a sequential stream, generating a sense of shared narrative trajectories – a sense of mutual occurrence or happening that unfolds, develops, reaches peaks or valleys, is periodically thrown off balance and must find resolution. In other words, nonverbal sonic fields entail a dynamic sense of motion, of a story being told, of “aboutness”. Although this narrative or motional sense can also be discerned in the speech of the single speaker, my focus here is on intersubjective shared narrative. But, since sonic signs are nondiscursive, the question arises - what kind of narrative it is?

In the last two decades, musical theory has addressed music from a narratological point of view\(^\text{113}\) (Abbate, 1991; Almén, 2008; Cone, 1974; Cumming, 1997; HaCohen, 2001a, 2001b; Hatten, 1994; Kramer, 2002; Monelle, 1992). According to Almén (2008), a musical narrative must embody an initial conflict, transgression, or opposition among elements; this produces a disequilibrium that becomes a source of dynamism for the unfolding process. Especially through semiotic concepts like

\(^{113}\) It must be stressed that the narratological approach to music does not address the question of program music. The narrative sense it addresses emerges through the semiotic operation of musical signs and not by pre-conceived discursive reference.
genre, style, markedness and virtual persona\textsuperscript{114}, as well as the explanation of phenomena such as tonal and motivic conflict, repetition, and resolution, these writers elucidate how sonic indexical and iconic signs create a sense of happening, cohering into structural entities who musically perform a “story”. These approaches are illuminating in that they demonstrate the possibility of conceiving a narrative power in nontexted music. However, because they deal with the specificity of music as an aesthetic language, and because they mostly engage in formal analyses of composed works of music, I consider them less suitable for my purposes.

The nature of the present object of investigation requires a narrative approach through which we can consider sonic occurrences that are not aesthetically constrained, but are, rather, spontaneous (improvised, taking place in the moment) occurring, \textit{in vivo}, between two interacting participants. Therefore, I find in the ideas of Levinson (2003) a better approach to musical narrative. In considering the experience of listening to music in the moment, he argues that music’s meaning is obtained non reflectively from listening to moment-to-moment occurrences. Here, the narrative aspect of music is not the same as the narratives of films or literature, where viewers and listeners can obtain a global sense of the plot the work unfolds, through characters who continue to develop its dramatic action. For Levinson, music can be considered narrative in a looser sense, which he calls kinetic-dynamic (a sense of “what’s coming next?”), as opposed to a global-spatial sense of architectonic narrative (“what’s the big picture?”). The kinetic-dynamic way of listening that Levinson describes is the type of narrative listening most relevant to our discussion.

\textsuperscript{114} These concepts are the outcome of the investigation of musical signification. In addition to the writers mentioned in this paragraph, the reader is referred to Chapter II.
For the purpose of understanding the narrative of the immediate vocal interaction, it may be enough that at each point one has an appropriate sense of an event arising out of and leading to another event, but without necessarily taking a synoptic view of what the series of events mean as a totality.

Coming from a different field altogether, though echoing Levinson’s approach, is Stern’s theory of the present moment as lived story (2004). His ideas both clarify Levinson’s idea of a kinetic-dynamic narrative in the unfolding of sonic events, and posit the conception of narrative in the intersubjective field. The phenomenon of a quasi-narrative occurring in the general nonverbal dimension, not only through the voice but also through gesture, posture and facial expression, has been addressed by Stern in great detail. He has named these short, present-time narratives “lived stories”, and characterized them as short emotional episodes, performed in the present moment. His emphasis is of course, on the intersubjective aspect of the narrative: emotionally charged intersubjective states are in fact, shared nonverbal lived stories. My focus is on the vocal aspect of those lived stories, to which I refer as shared vocal voyages.

Stern (2004) claims that a present moment\(^\text{116}\) contains the essential elements that compose a lived story (it is lived as it happens, not as is put into words afterwards). It is largely nonverbal, and, most importantly, it is of very short duration compared to most stories (roughly a few seconds). Made up mostly of emotions that unfold in

\(^{116}\) Stern’s view of a present moment (parallel to the concept of “psychological presence” in cognitive theories) is comprised of a sense of immediate presence, of an approximate length of 4-8 seconds. For an in depth explanation of this phenomenological conception, see Stern, (2004), part I.
the moment as an untold emotional narrative, the lived story contains elements that are common to the verbally told story: a trigger for drama that needs resolution, a plot, and a line of dramatic tension that pushes the story forward from the buildup, through crisis and toward resolution. This ties the story and its composing elements together temporally (or as Levinson phrases it, creates the concatenation of events).  

The next stage that interests us is that the present moment as a lived story can be shared. It is in these shared lived stories as they occur in the vocal sphere (hence, vocal shared voyages) that I am interested here. Stern claims that when lived stories are shared in the moment, intersubjectivity begins to take on flesh. Significantly, he views these lived stories in the moment as building blocks for larger narrative (cognitive and verbal) structures. The small, nonverbal narratives that unfold in the present moment are nested and embedded in the large narrative that unfolds over many clinical meetings. In fact, Stern points out that the clinical encounter works in a recursive manner, as the small moment to moment units share a similar narrative structure with the larger historical one. Stern's idea of the present moment as “a world in a grain of sand” contains the understanding that behaviors in the present moment largely represent patterns of behavior that have crystallized in the history of the individual, reflecting patterns of intersubjective relations experienced and

---

116 The lived story was initially termed as a pre-narrative envelope (or proto-narrative envelope) by Stern (1992). However, since pre or proto imply primitiveness and since he views this phenomenon as “not primitive at all, but rather as a fully developed, normal and pervasive aspect of children’s and adults’ lives”; an emotional narrative that is felt rather than cognitively and verbally constructed, he shifted the name to lived story. (Stern, 2004, p. 58).
learned in the past. Stern (and from a musical angle, Cumming, 2000) suggests that a close attention to the performance of an individual in the present moment (as opposed to the traditional approach where the past is told verbally by the patient), can disclose the past, in terms of intersubjective patterns, expectations and attitudes: the wordless narrative of the present moment contains within it the larger historical narrative of the individual. Killingmo (1990), for example, describes the deportment of a female patient as follows:

Her use of words, as long as it takes place in written form, seems to be comfortable and free of conflict. This state of affairs is radically changed when she tries to express herself verbally, that is to give sound to words. Her speech is high and unmodulated, giving an impression of impartiality and lack of concern. Also, her general appearance expresses a sort of virginity. Lacking the sensuality of an adult woman she reminds one of the untouched daughter of a mighty freeholder from Norwegian folktales, a long-legged chit of a girl on the edge of puberty. (Killingmo, 1990, p. 114).

In the course of her treatment it gradually became clear that the patient’s relations with words and language were disturbed by the defense mechanisms of intellectualization and isolation:

---

117 The philosophic idea that the past is folded in the present is of course not Stern’s original idea. Husserl (1964) and Bergson (2001/1913) are two of its main proponents. In the present context, this also resonates with Cumming’s (2000) idea that performance in the present time can be understood as an expression of identity (which obviously has evolved over a long period of time). The condensation of past and future in the present of the musical experience is one of its main features and has been addressed by several music authors (for example Meyer, 1956, Zuckerkandl, 1957). Rose (2004, pp?) provides a clinical example that demonstrates this temporal condensation as occurring in both music and the clinical encounter.
Among other things this became visible when she took one of her characteristic pauses—carefully searching for the appropriate word. This hesitation was due not only to a fear of choosing a word that might be too emphatic and thereby running the risk of being offensive or opinionated. Another fear was also present, that of putting words to use at all. She behaved as if a word in itself could be a dangerous thing and one cannot know what may happen if one takes it in one's mouth. Therefore the sound itself had to be isolated from the very image of the word. The result was that certain words were never employed at all in her conversation.

The peculiarities of the patient’s speech were subsequently related to her relationship with her mother, an utterly restrained woman that exhibited extreme anxiety concerning any kind of emotional excitation, allowing no spontaneous liveliness or outburst of temper. In a very subtle manner, the mother continually disqualified the very being of her daughter, especially by disregarding and by scorning her distinctive, lively character.

This left the patient in a panic that mother might leave her. Thus the originally impulsive child became an extremely obedient and nice little girl as seen from the outside. Every experience related to affect was located within the world of fantasy sharply isolated from the world of reality. In fantasy she could indulge in excitation. At the same time reality was drained of emotionality, remaining grey and devoid of attraction. To manage this double life, her inner world had to be kept secret. Therefore words, too, had
to be stripped of all emotionality before leaving the mouth, ensuring that nobody could take her at her word. (ibid., pp. 120-121).

In a recent paper, Gratier and Trevarthen (2008) investigate nonverbal narrative as it develops in nonverbal communication between mothers and infants. Through a rapidly growing sense of rhythm and control over their movements and vocal production, infants begin to express their sense of engagement as well as their intersubjective intentions vis a vis meaningful others. These proto-conversations, or narratives without words, unfold between mothers and infants essentially through live performance (and not through linguistic construction of themes and plot lines). Gratier and Trevarthen maintain that

in live narrative, it is the musicality of voice and body that holds listeners’ attention so that they can be brought simultaneously to share in the experience of a dramatic, challenging moment... Written narrative which came out of oral narrative may be seen as involving not only cognitive processes of understanding and remembering but also embodied perceptions of the tension and drama of shared emotion ... (2008, p. 124).

The main idea that emerges here is that the narrative of the nonlinguistic voice constitutes an additional level of communication, one which has vital meanings and which unfolds simultaneously with verbal communication. One of its important meanings is related to the fact that this communication is co-constructed in the moment. The sense of “happening” in a conversation is an outcome not only of its referential contents but of its performance. The lived stories that are spun in the
moment are, in fact, events\textsuperscript{118} – vital and unique happenings between us in the Now. Moreover, it is not just happening, it is we who are creating the happening as we go along. This sense of mutual participation in a performance in the Now is central to the idea of vocal narrative.

But what about content? What is the narrative about? In line with Gratier, Levinson, Stern and Trevarthen, I maintain that the shared vocal voyage is self-referential – it is about the partners' ongoing relationship and communication, about what kind of experience it is for each partner to make the voyage together, about how the partners create their togetherness, about their expectations and boundaries, how they resolve the problems that arise on their way through the verbal or nonverbal plane. In this light, Killingmo's clinical example reveals how the vocal behavior of the patient indicated a serious early disturbance of affective communication between her and her mother. As a consequence of this understanding, Killingmo describes his own modified interaction with the patient, whereby he abstained from relying on semantically based interpretations while conveying attentiveness and affirmation through his own vocal intonation.

Beebe et al. (1997) similarly argue that in mother-infant interaction, what is represented presymbolically is the interaction process itself. For example, new ideas brought verbally to the interaction present the partners with new inputs that arouse emotional reactions. These are processed not only intellectually, but through performance – and heard through vocal and rhythmic display. At times, problems are presented nonverbally - for example, by a display of difference or a mismatch in

\textsuperscript{118} Recall Bakhtin's idea of utterance as an event, in other words, as a unique happening taking place in each moment.
tempo or tone, by the changing quality of mutual attention, etc. Resolving such problems and differences affects the implicit therapeutic relationship and the sense of alliance. In the course of therapy, this sense is strengthened or weakened according to the intersubjective experiences accumulated through the concatenation of these shared vocal voyages (see the concept of “now moments” above).

**A Short Note on Rhythm**

Time and rhythm are fundamental elements both in music and in human experience. It would be expected that a study of the voice would devote a key section to time – as the natural medium of sound and music. However, the magnitude of the subject is such that a thorough discussion is beyond the scope of the present study. In addition, time and its vast significations (physical, musical, philosophical, subjective and metaphoric) have been abundantly addressed and debated throughout history within the disciplines of music, philosophy, psychoanalysis and development.\(^{119}\)

What is of importance for our purposes is the function and signification of time within the vocal phenomena I describe below. Temporal organization underlies and permeates all vocal forms, and I will address this aspect throughout the next section.

At present, I will state, however, that rhythm in sound is linked to a complex field where several variables are at play (for example pitch, duration, volume, pause). Temporal organization of these variables occurs at several levels of the sonic field, and it is important for our purposes to distinguish between three such organizations.

\(^{119}\) Much literature on the various issues comprising the subject of time also forms part of the present research. See for example Bergson, 2001/1913; Husserl, (1964); Langer, (1953); Meissner, (2007); Rose (2004), Stern (2004), Zuckerkandl (1957).
1. Beat: The first fundamental temporal organization is the phenomenon of pulse, the basic, recurring time unit of music, also known as “beat”. In popular use, beat often refers to the general tempo of a piece, the meter, or “groove” (see p. 88 below). Most music is characterized by recurring cycles of stressed and unstressed beats (often called “strong” and “weak”) organized into measures (short groupings of patterns or beats), where phenomena are grouped into similar clusters that recur in relative cyclicity. It has been demonstrated that, as of the beginning of life, individual speech and dyadic vocal interaction are organized by a recurring pulse, though less precisely than in music (Mazokopaki and Kugiumutzakis, 2009).

2. Rhythmic pattern: Occurring simultaneously with beat, a pattern is created when a certain rhythmic grouping of sounds, pauses, stresses and durations are organized into a short, recurring temporal form. Repetition of rhythmic patterns create an additional sense of time to that of beat, by creating instances of stress other than the beat. Rhythmic patterns in speech are similar to poetic meter forms, and are perceived in utterance through vocal gestures (see below). Rhythmic patterns can occur in correspondence with verbal production (this would be the prosodic aspect of speech). However, as I maintain, they can be organized within speech disregarding the semantic level, as a repeating musical form that underlies linguistic performance. The meaning of this phenomenon will be elucidated in the section on vocal gesture below.

3. Tempo: This is the general speed or pace of a sonic section, whether short or long. A speaker's vocal gestures (having a distinct rhythmic pattern) can be
performed at a greater or lesser tempo. Tempo can characterize the general flow of an individual speaker or of dyadic interaction or only parts of these.

The three temporal organizations of beat (pulse), rhythmic pattern and tempo can occur in speech with varying degrees of salience. Vocal expression may appear to be lacking beat or rhythm, or conversely, as having a pronounced rhythmic feel. Since rhythmic organization is fundamental to music, language and all movement and cultural practice, these temporal forms are powerful indices and icons of subjective and intersubjective emotional states. The principle of correspondence described above is principally realized through the co-matching of temporal elements. Feelings of lack of correspondence or togetherness in interaction are often characterized by a lack of temporal correspondence (quintessentially, when the participants fail to find a common beat – as when one of the speakers displays a tempo that is markedly different than the other speaker, or when the partners do not match the length of the switching pause\textsuperscript{120}).

III.4.3 Vocal Forms and Phenomena

Vocal timbre

Vocal tone is referred to quite often in the analytic literature to denote the quality of the speaker’s voice. The phenomenal experience of the other’s voice is implicitly read as an index of the speaker’s emotional or energetic states (as when someone’s voice is heard as sad, irritated, tired or excited, allowing the listener to infer the

\textsuperscript{120} The concept of the switching pause is explained in the section on Pause below.
speaker’s situation) and/or as an icon of the speaker's personality (the speaker’s vocal tone sonically portrays something more stable, that is felt to reflect her character or identity). In the last decades, vocal tone features in the literature as one of the main components of the nonverbal dimension. More often than not, analytic writers demonstrate a correct intuition about the meaning of vocal tone and the importance of listening to it. Nevertheless, in the following elaboration I want to highlight its importance to the intersubjective dimension.

There is a certain ambiguity in the literature in the use of the term “vocal tone”. It is often unclear whether writers refer to vocal color, or vocal intonation (i.e. vocal melodic contours, or inflection)\(^1\). In order to distinguish between the two, I will here use the musical concept of *timbre*, which is one of sound's basic components\(^2\), and unequivocally denotes the *quality* of the voice’s sound. First, a physical explanation: Most oscillators, from the human voice to a guitar string to a bell, simultaneously vibrate at a series of distinct tones having different frequencies. The lowest normal mode frequency is known as the fundamental frequency, while the higher frequencies are called overtones and are more difficult to perceive aurally. Timbre is the quality that enables the listener to distinguish between the sound of different instruments or voices.

Vocal timbre is physiologically created in the larynx as a specific combination of overtones that resonates when air coming from the lungs creates friction against the vocal folds, so that every sound one makes represents a specific combination of

---

1. The confusion is understandable, since in order to discern vocal color, sound must have some temporal duration, which means that it is blended into other vocal occurrences, like intonation and gesture, making it difficult to tell these phenomena apart.
2. The etymology of “timbre” is from the old French in which it means “bell”.  

178
overtones. The timbre of a sound is determined by which overtones it emphasizes. That is to say, the volumes of these overtones, as related to each other, determine the specific color or texture of any sound. Like cloth, rope, or wood, whose texture is determined by the materiality of the fibers composing them, the specific combination of vocal “fibers” (overtones) produces aural texture.\textsuperscript{123} And just as our touch recognizes – in a bodily, non-reflective mode of cognition – the texture of a surface, so is the ear sensorially impressed by the speaker’s vocal timbre.\textsuperscript{124}

Vocal timbre is not stable: the intensity of each of these overtones is rarely constant for the duration of a sound. Over time, different overtones may decay at different rates, causing the relative intensity of each overtone to rise or fall independently of the overall volume of the sound. In addition, other elements like sound onset (called in music “attack”), which is the initial contact or friction between the vocal folds and the air, as well as the intensity, the vocal range and the lability of the speaker’s intonation all affect vocal timbre. These elements, no less than the average sonic frequency of the voice, are active in the creation of the speaker’s characteristic vocal timbre. A careful listening can detect timbral changes in the moment. The hypervigilance of traumatized persons to changes in the interlocutor’s vocal tone is a well known phenomenon. Although vocal timbre is dynamic and labile, it is nevertheless phenomenally perceived as a global characteristic of a speaker, a token of the speaker’s identity, or an “acoustic fingerprint” (Dolar, 2007).

\textsuperscript{123} Roland Barthes (1977), famously referred to it as “the grain of the voice”.
\textsuperscript{124} This of course includes the speaker’s perception of her own vocal timbre. The analyst’s vocal self-impression features rarely in the literature: it is ordinarily the patient’s vocal timbre that is noticed. This situation has much changed in the last decades with the development of intersubjective approaches to psychoanalysis. Interestingly, in literature about psychological treatment written by patients, the vocal qualities of the analyst are frequently referred to (see for example Chernin, 1996).
Obviously, timbre is affected not only by physiological factors but also by emotional, interpersonal and cultural ones. For example, the voice is acutely gendered, and its feminine and masculine manifestations are mostly designed according to cultural expectation and norm. Thus, a female's voice is expected to sound soft and high-pitched, while a male's voice - more roughly textured and deep. A quality of roughness in a female's voice (and vice versa, feminine attributes in a male’s voice), arouses special attention and emotional response. The same can be said about other timbral qualities that come to signify age, social status, national identity, etc. All these create invariant timbral qualities that have hermeneutic value, as they form identities, and give rise to expectations and emotional attitudes.

These considerations are certainly important, though they are not the center of my present concern. What I propose in the following, is to consider vocal timbre’s subjective and intersubjective value in the present moment of the clinical encounter. As explained in the introduction to this section, nonverbal vocal phenomena function simultaneously as “articulators” (by actively giving form to the sonic stream) and as “signifiers” (by giving feedback and information as to what is happening), hence:

As articulator, vocal timbre works in an impressionistic way. It affects and influences the intersubjective vocal field by impressing listeners, actively imparting a sense of texture, or “feel” to the vocal exchange, thus arousing affect and response. As signifier, vocal timbre acts as a “sonic sign of subjectivity” (to borrow Cumming’s [2000] apt phrasing) and, I hasten to add, of intersubjectivity. Vocal timbre is a sign (both indexical and iconic) of the speakers’ felt, subjective and intersubjective experience in the moment.
In order to elucidate these ideas, let us first consider the fact that, unlike other sonic parameters like pitch (that has proper names – note names, or volume than can be partially quantified and qualified through natural linguistic markers), vocal timbre has no linguistic labels of its own. Adjectives, adverbs and metaphors are needed to describe its infinite nuances. These corporeal metaphors are sometimes energetic (a forceful tone, a lively tone, a dead tone), sometimes coming from the sensory realm (like the visual: a clear, bright or dark voice), are the subject of an emotional qualification (a loving, anxious, or arrogant tone of voice), or of a spatial one (a penetrating tone of voice, a distant tone of voice).

Here, however, I am particularly interested in tactile metaphors that qualify the voice as smooth, rough, warm, cold, soft or hard. These metaphors indicate that aural-vocal interaction is often experienced as a form of touch. Indeed, through vocal exchange conversational partners touch each other and interpenetrate. This phenomenal experience can be explained physically, since sound waves do physically reach, or “touch” the eardrum. But this is not the key explanation. Rather, through the capacity for cross-modal perception, the voice’s aural global qualities can be perceived in terms of another sense, in this case – the sense of touch.

The tactile metaphors used to describe vocal timbre disclose the inherent intersubjective nature of the voice. When it appears in its corporeality in the shared auditory sphere (and in the analytic situation where almost every corporeal interaction is suspended, voice represents the only form of touch possible), vocal timbre evokes affective responses. In this manner, it acts as the articulator of the intersubjective field: it reaches out to the other, making contact with him by
generating powerful feelings like curiosity or indifference, sympathy, attraction, rejection, pleasure or abjection. Intersubjectively, vocal timbre is often experienced (though prereflectively) as exciting, soothing, seductive or violent, as provoking anxiety, eliciting closeness or distance, and driving the listener, often without his awareness, to respond.

This means that, as listeners, we are not objective bystanders, simply taking in the speaker's vocal timbre and decoding it: in a prereflective manner, we enter into a relation with it, we are touched, activated by it and respond to it. Our tactile reading of the voice is intersubjective, and the question can be asked: How does your voice touch me, move me, arouse me, and how do I respond? Conversely, when we hear the timbre of our own voice in the course of the intersubjective encounter, we sense not only whether its tone is alive, dynamic or rigid, whether it resonates or is blocked, but also how it reaches out to our listening partner, how it touches, arouses, envelopes, calms or excites the other.

It is important to note that, to no small extent, vocal timbre can be controlled and manipulated, and analytic partners do so in order to relate and correspond, by reading, negotiating, matching or mismatching their vocal timbre. The dramatic therapeutic effects of the manipulation of vocal timbre are described in the analytic literature more than once (see for example Killingmo above; Ogden, 1999a).

Regarding the historical intersubjective sources of vocal timbre, I argue that voice and listening form and constitute each other: A person's general timbre as her sounding identity, or acoustic presentation in the interpersonal world begins through early, concrete intersubjective vocal-auditory contact. The Romanian violinist Kató
Havas said that she could tell, when listening to a violin’s sound, how it was handled by its player: whether it was played with respect and sensitivity, whether the violinist developed and cultivated the responsiveness and resonance of the violin’s wood, or whether it was played insensitively, leaving its sound opaque and sealed. Thus, the violin’s sound carries the signs of its player (reported in Cumming, 2000, pp. 127-128). The voice of the speaker carries similar traces. Vocal timbre reflects the historic, intersubjective vocal contact between mother and baby and their mutual listening: How did they vocally touch each other? Was it a pleasurable, musical, reciprocal touch, a touch that stimulated the baby to resound and feel pleasure in her acoustic presence, or did she learn to turn down her voice and play down its resonance? Of course, we must also ask how the mother is touched by the baby’s vocal tone, and by analogy, how is the analyst touched by the analysand’s vocal timbre (this focus on the caretaker’s subjectivity is central in the thinking of the psychoanalyst Jessica Benjamin [1988] however the vocal manifestation of this issue is not addressed).

Vocal timbre carries a history of reception, developed through the presence of the mother, who was not only an acoustic container but also a primordial audience. The mother’s responsiveness to the baby's voice and its quality acts upon the baby like the acoustic qualities of a music hall and its audience act upon a performing musician. The speaker’s vocal timbre tells the story of the emotional resonance of the listening mother – her pleasure, anxiety and fantasy, her expectations and wishes. Thus, in the intersubjective vocal contact that builds up throughout the analytic process, vocal timbre can be understood as a phenomenon that has history,
that carries within it previous addressees and contains significant others. Through listening to the effects of vocal timbre, hypotheses can be conjectured about intersubjective historic environments, implicit modes of addressivity, the vocal creation of psychic defenses, and so forth.

* * *

I have discussed the category of vocal timbre in some length, in order to exemplify the way I listen to all nonlinguistic vocal phenomena. As can be seen, this first entails a phenomenal impressionistic listening (phenomenologically speaking, a consciously “naïve” approach and consequent description of the phenomenon). Secondly, an understanding emerges as to the signifying power of phenomena both in terms of the uttering individual and as a reflection of the dyadic field, that includes present and past dyadic patterns brought into the present situation by both partners. The following vocal phenomena will be more briefly addressed, but the same interpretive process can be applied to all.

**Vocal Gesture**

One of the phenomena that emerge out of the stream of vocal communication is vocal gesture. A single vocal gesture is a short unit that has the quality of a melodic motif. Neuropsychologist and pianist Manfred Clynes defines a gesture as a short motion that can be perceived and measured: having a shaped contour, with varying direction, levels of pressure and emphasis and overall length (a gesture is phenomenally felt to be increasing or decreasing, going “upwards”, “downwards”, penetrating or fading out, etc.). He identified a set of basic gestural motions that are
neurophysiologically encoded to express basic emotional states: sex (erotic love),
anger, platonic love, grief, joy, hate and reverence (in Cumming, 2000, p. 140; Lidov,
2005, and also Stern, 1985). Clynes argued that such gestures could be carried out
by any part of the body – their dynamic forms could be expressed through any motor
output, including, of course, the vocal apparatus.

Vocal gestures display a more complex articulation than timbre and are thus higher
in the hierarchy of vocal forms (in turn, gestures compose more complex structures
like vocal persona). Formed out of several sonic components (pitch, timbre,
rhythmic pattern, tempo, emphasis), they have a melodious inflection that imparts
motion and vitality to speech. While they are integrally fused in the total utterance
(they are not reflectively superimposed on words)¹²⁵ vocal gestures generate
meanings and create effects that are partially independent from verbal semantics.
These cannot be annotated but can be acutely sensed. Listeners intuitively identify
and sense the contour – onset, duration and end – of a vocal gesture, even if it is
minimal, and prereflectively interpret its meaning.

Gestures are addressive in a more distinct and focused manner than timbre – they
are more clearly directed at an interlocutor and more eloquently articulate affective
and communicational intents, hence imparting more information and more directly
eliciting the listener’s response and participation. Like all nonlinguistic vocal
phenomena, gestures are part of conventional vocal dialects - for example agreed-
upon vocal inflections convey irony, joy, ambivalence, etc. But vocal gestures can
also be idiolectic, conveying uniquely personal, idiosyncratic meanings. In addition,

¹²⁵ On the other hand, it is precisely these aspects that are manipulated in the performing arts –
theater, dance and music.
vocal gestures can appear independently from words altogether as nonverbal utterances (like “uh”, “hmmm”, etc.).

The subject of vocal gesture, although ubiquitous in speech, is not very prevalent in analytic literature. Though it is sometimes acknowledged as an important nonverbal sign, it is usually mentioned in passing, and even then, implicitly so. In recent decades it features more often, as a result of an increased interest in the nonverbal dimension of analytic communication, but, as in the case of timbre, a certain vagueness about its precise meaning arises and some clarification is required regarding its sonic structure and power of signification. Music theory has in the last decades widely addressed musical gestures, viewing them as a central musical sign and compositional device. Although this aesthetic discourse obviously differs from the psychoanalytical, it can nevertheless contribute to the understanding of vocal gestures in the clinical setting.

The music semiotician David Lidov (2005) posits that the general significance of music emerges from its links with somatic experience: music bears meaning only if its capacity to convey somatic experience is understood – although, obviously, the connection between the body and musical expression is neither direct nor simple. Within musical discourse, this position represents a radical departure from formal approaches to music. We do not need to consider these issues here, but for our present purposes, the equivalence that Lidov creates \(^{126}\) between musical figures and bodily experience is constructive.

\(^{126}\) As well as other music theorists like Clynes, Coker, Tarasti, and Hatten (in Lidov, 2005, especially chapter 8).
Lidov defines musical gestures as brief, expressive, molar (i.e. whole, indivisible) melodic patterns that are indices and icons of motor activity or somatic experience. These units are perceived, in a gestalt manner, as sonic envelopes\textsuperscript{127} that signify diverse states like halting pathos, choking grief, melancholy, dancing, leaping, grace, mounting excitement or tension, shyness, hesitancy, etc. Lidov addresses the aesthetic question regarding the manner in which musical figures “translate” muscular gestures. He postulates a certain isomorphism between musical and muscular gestures, suggesting that they are linked by an inner behavior correlated with a state or feeling. Obviously, there is a structural difference between the two: bodily gestures are continuous motions, while musical gestures are composed of distinct, different notes – though they are perceived as one sweeping motion (this is one of music’s central illusory effects). Lidov assumes that “the notion of expressive gesture interpreted as the surface form of an underlying neurological function gives a logical connection between somatic and musical experience, one which shows a basis for cause and resemblance in the relation of music to feeling” (2005, p. 151. See also Webb, 1760/2003).

It is precisely this connection that has been elaborated and empirically demonstrated by infant researchers. As documented by the growing body of developmental research, dyadic communication develops in the first weeks of life.

\textsuperscript{127} The musical discussion of gesture makes a distinction between gestures as compositional entities notated in musical scores and as inflected gestures performed by instrumentalists (see Cumming, 2000). Indirectly, this issue is relevant to the analytic situation: here we are dealing exclusively with performed vocal gestures that cannot be notated. The prevalent analytic custom of writing down verbal protocols of analytic sessions, necessarily omitting the registration of vocal gestures, raises serious questions about such reporting methods as they inevitably exclude important dimensions of information from analytical consciousness and discourse.
through vocal and bodily gestures, expressive of the partners' intersubjective motives. When mothers and babies communicate effectively, it is clear that they are highly attuned to their mutual bodily and vocal gestures. Without this attunement, there develops a severe disruption of the self and of intersubjective capabilities. As we have seen above, Stern’s concepts of vitality affects and amodal perception facilitate our understanding of how certain attributes of experience can be abstracted from one sense domain (for example proprioception), into another (for example vocal expression). Thus, the translation, or transaction, between, say, an inner sensation or emotion (like physical tension or exuberance), and their vocal expression, although not yet completely explained by science, is demonstrated empirically to exist in newborns as an innate ability.

As in music, the main function of gestures in psychoanalysis is communicative (by both conveying information about the speaker’s state and having an addressive impulse, eliciting response from the listener). Perception and interpretation of vocal gestures occurs largely out of awareness through processes of somatic identification. But, differently than in the musical field, where gestures are highly stylized aesthetic forms, thus indirect and distanced from simplistic somatic expressions (although, on the other hand, perhaps more appealing, because of their clarity and specificity), vocal gestures in speech are more immediately related to speakers’ here-and-now states. It will be remembered, however, that vocal expression, too, is in no small measure stylized and manipulated. Thus, vocal gestures contain both a directness (through this they are indexical, relating to the immediacy of the moment), and a more deferred manipulation (making them more iconic, i.e. organized and stylized,
though not for artistic purposes and mostly out of awareness, signifying something about the speaker’s identity, choices and intersubjective history).\textsuperscript{128}

In terms of the individual, vocal gestures are “signs of subjectivity” (Cumming, 2000), of the particular speaker. As signs, vocal gestures signify in indexical and iconic ways: they reflect stable or fleeting subjective states of vitality and affect in the speaker. For example, a hesitating vocal gesture may be part of a local (or permanent) state of hesitancy, characterized by an arrest in the flow of energetic pace and held ambivalently in a state of suspension. Vocally it can take many nuanced forms, but usually it will include elements of temporal arrest or delay, pause, lowering of volume, suspension of sound, or the stressing or prolongation of sounds. The semiotics of hesitancy creates a certain music in speech. As Merleau-Ponty suggests (1962/2002), the hesitating vocal gesture – like all other energetic and affective gestures, are intuitively understood by listeners because their shape is known to us through our own (embodied) hesitating experiences (of delay, arrest of energy, sense of suspension, etc.). This conception resonates with such isomorphic approaches as those of Lidov, which we have seen above, as well as Langer (1970).

On the other hand, vocal gestures represent the subject’s choice of articulation (whether this choice is conscious or not). If repeated, a vocal gesture becomes a general feature of the speaker’s deportment and a building block of her vocal persona: a perceived aggressive character, for example, may repeatedly perform

\textsuperscript{128}In addition to their somatic signification, vocal gestures often refer to parallel verbal semantics – the effect of irony, as a special polyphony between words and gestures, is such a case. However, this phenomenon typically belongs to the domain of paralinguistics, and will not be dealt with here.
Vocal gestures with characteristic inflected contours that are unpleasantly experienced as intrusive or forceful (more on this below).

Vocal gestures are at times more local occurrences, characterizing specific themes. Patients often display specific vocal gestures when speaking of specific self-experiences. One patient would use rapid, energetic vocal gestures when speaking of her professional self. In this case, the marked velocity and specific stress in her gestures created a sonic icon of self assurance and assertiveness. In contrast, when talking about her parental role, her gestures would drastically change and create, through temporal delay and weak stress, a vocal icon of hesitancy and weakness. The different “selves” she experienced through these two differing roles were more markedly manifest in her vocal gestures than through her verbal discourse. The analytic notion of the multiple selves that comprise selfhood has been suggested in the last decades by relational analysts - for example by Bromberg (1994) and Mitchell (1993). I suggest that these multiple selves can often be perceived through distinct vocal manifestations (this theme is further addressed in the section on vocal persona).

Vocal gestures also characterize specific themes and figures. Whenever he talked about his father, a patient would consistently perform sneering gestures through a heavy nasal timbre, creating an effect of ridicule. Bringing this to awareness facilitated a discussion of his disavowed feelings of anger towards his father.

---

129 Clinical examples are provided both from the literature and from my own clinical experience as therapist.
Gestures may also appear as single occurrences, creating dramatic effects of surprise or contrast. Ringstrom (2001) gives a humorous example of his own vocal gesture that worked sharply as an interpretive act of Jonathan’s (his patient) recurring self deprecation. The exact inflection of the analyst’s vocal gesture can only be imagined through the written text (and of course, different speakers would perform it differently):

I started to feel trapped by Jonathan's vitriolic self-abuse. Although this old transference–countertransference script was doubtlessly familiar to me, I blurted out, 'But of course, Jonathan, your stupidity was sooo understandable!' He howled with a laughter more intense than I could remember witnessing; he also looked relieved. He immediately repeated my words three times—tone and inflection—in a state of astonishment and childlike glee. 'But of course!' he said, 'My stupidity was sooo understandable!' Before the end of the session, he repeated the comment several more times with great appreciation and expressed his wish to remember it exactly as I had said it. (Ringstrom, 2001, p. 728, italics added)

This example demonstrates the interpretive power of performative acts – such as gestures. What was significant here was the ironic inflection in vocalizing the word “so”. Similarly, Stern (1998) discussed the meaning of heightened affective moments (“now moments”) occurring in therapy largely through nonverbal communication. This example also shows that, although gestures are ordinarily
perceived intuitively and out of awareness, they can at times become conscious and be used therapeutically.

As for their articulative intersubjective role, vocal gestures function as relational acts. Beyond their possible individual content, vocal gestures are acts of address. They appeal to – one could even say compel - the listener’s attention; pleading, cajoling, rebuffing, warning, ordering or apologizing. Vocal gestures entangle the listener and demand her interpretation and response, primarily through sympathetic, prreflctive means, often through the mechanisms of matching and mismatching, or at times through a quasi-musical mechanism of call-and-response. Vocal gestures can be heard as micro-narratives because they have a certain temporal-dynamic trajectory. As I will demonstrate in the section on vocal duet, it is possible to listen to the clinical dialogue as one listens to an improvised musical duet, containing dynamic, relational vocal gestures that create a relational ebb and flow, generating dyadic effects of sympathetic resonance, “moving along” together\(^{130}\), processes of coming close or pulling away, domination or giving in, etc. Finally, as is the case with all vocal phenomena, the intersubjective value of vocal gesture includes not only the present analytic partner, but historical addressees as well. Thus, vocal gestures can be partially understood as reverberating past vocal interactions and relational patterns (see an example in the section on vocal persona).

\(^{130}\) “Moving along” is Stern’s (1985, 2004) important concept of proceeding through a therapy session on a local, moment-to-moment level. It is an unknown, unpredictable process consisting of relational moves that strung together make up the totality of the session.
Silence and Pause

Silence in therapy - something that seems like nothing, a no-thing that is a thing\(^{131}\) - presents participants with a paradox. For psychoanalysis, silence had initially been a problem. After all, it was termed “the talking cure”, and verbal expression was the phenomenon under inquiry. Psychoanalysis investigated words and language and, through them, made claims about the existence of an unconscious. For a while, silence was apprehended as resistance to the analytic process, representing hostile or aggressive feelings towards the analyst. And indeed it may be so. But it was Winnicott (1963/1984b, 1989), Balint (1968), and others who, staking a claim for the patient’s “right to not communicate” (Winnicott, ibid.), allowed silence to be treated more benevolently, as a defense in the transference which must be listened to sensitively and with more acceptance. Silence brought to the forefront the idea of the analytic encounter as an experience in itself: being-with the patient (not only attending to his verbal associations) began to acquire new meanings, and be considered as a major therapeutic experience. Silence began to be experienced and listened to as a material in itself, not as no-thing. From a developmental point of view, Stern (1985) seems to support this direction, by indicating the losses that the learning of language brings about - words can only carve out parts of subjective experience into the social, declared domain. As he argues, the rest recedes into silence: the realm of the personal, the idiosyncratic, the unspoken, very often remaining unvalidated by others and ultimately denied to the self. Silence in therapy – in other words, its appearance in the presence of a listener - is a special

\(^{131}\) It is Bion’s illuminative thinking that articulated the idea of “thing” and “no-thing” in relation to silence (1957, see especially chapter 12).
opportunity in which the subject’s most private, unspoken existence may be known and shared, and perhaps even given a voice.

It is interesting to note that even within this new analytic attitude, most of the literature on silence deals with what I call “major” silence in analysis. Silence came to the focus of attention when it appeared “in bulk”, so to speak, in patients who kept silent for long stretches of time, for whole sessions and sometimes even for months (see Coltart, 1993a, 1993b). But the fact is that silence, as a soundless phenomenon, is a primary material, permeating the dialogue through and through. Not only does it segment the stream of sound into phonemes, words and phrases, but, materializing in myriad forms of pauses and breaks of varying duration and quality, it is central in structuring the dialogue’s rhythm and articulating its musical flow, as well as being an utterance in its own right. It is to these manifestations of silence which I would now like to draw attention: the pauses, cues, delays, breaks, gaps, gasps, breathing, resting and transition points that are part and parcel of the dialogue. Although they are mentioned in clinical texts often enough, they are mostly unaccounted for, so intrinsic and implicit they are in communication (microanalytic approaches to analytic interaction are an important exception).

Let us consider the pause – the “small silence”. The pause is described often in the clinical literature, but only implicitly (an outstanding exception is Bion, 1957).

However, the pause has recently been addressed by developmentalists due to their

132 Conceptually, it may have proven easier to package those large chunks of silence into a discrete topic – “the silent patient”. It is also possible that, though difficult to sustain, the silent situation does not radically affect the analyst’s basic (silent) listening stance. But as we shall see, other forms of silence force the analyst out of this professional stand and forcibly cause him to participate more actively.
close attention to the micro-level of communication. Beebe, Alson, Jaffe, Feldstein & Crown (see also Beebe, 2005) described the phenomenon of the “switching pause” as part of the structure of social conversation that occurs through turn-taking. Mothers and infants pause for similar duration of time (i.e. they match the length of the pause) before the other takes a turn. Thus, switching pauses mark the boundaries of the turn exchange. Matching the switching pause means regulating the exchange of turns at a similar pace, or tempo (and well-matched switching pauses delineate characteristic rhythmic patterns in the dialogue, strengthening the sense of correspondence). Switching pauses are similarly matched in adult conversation as well. Partners tacitly know whether pauses occurring during the speaker’s utterances indicate that she intends to resume talking after the pause (this is an intrapersonal pause), – or whether they signal the end of the speaker’s turn, where the listener switches roles and begins her turn as speaker. These are interpersonal pauses, because they mark the boundary of the switch from one speaker to the other.

Beebe (ibid.) emphasizes the subtle interpersonal processes that occur in the matching of the switching pause. These processes are performative and implicit, occurring out of awareness. However, as we have seen above, vocal events come to represent emotional relational meanings. As some analytic writers have succinctly described it, a short pause may contain an entire drama, manifesting worlds of anguish, hostility, misunderstanding, loneliness, mourning, intimacy or erotic tension. And yet, a pause can be a place for rest, reflection and ordinary going-on-being.
A successful matching of the switching pause underlies the partners' feelings of good coordination, togetherness and being-with (in popular musical terms we would say a sense of “grooving”). When the normative length of the switching pause is altered, either by prolonging or shortening it, attention is inevitably directed to the pause itself – it stops being implicit, and becomes self-referential and conscious to the partners. And, as is the case with all vocal presentational signs, it is here that the experiential, performative factor (rather than the referential) comes into action.

Marking a change from the expected, the altered pause represents a new event in the exchange: in sonic terms, a space or gap is created, that needs to be renegotiated, i.e. interpreted by partners in order to establish both its possible meaning and what decisions are to be taken in order to continue moving along (when should speakers resume the talking, how, and by whom?). In music – whether pre-composed or improvised - pauses and short time gaps are fundamental rhetoric devices, central in building up expressive and dramatic effects. These small silences are both deliberate and stylized and maintain idiomatic propriety.

Throughout their training, performing musicians learn to attend to, as well as create pauses. Listening to their expressiveness in the analytic dialogue requires, perhaps, a musical ear of sorts. That is to say, in parallel to a cognizant linguistic listening, attention must be paid to the form and concatenation of sounds and silences. In this manner, pause as performance and interpersonal event should be heard in context, and its communicative value understood. Since the pause as sign is not discursive, it cannot be initially understood in a linguistic sense, but only lived through, experienced, much like the way improvising musicians attend to and interpret pauses and take decisions regarding them. What is important here is not so much
their possible specific meanings – since pauses can mean anything - but rather their power as an intersubjective event and a dialogic impulse.

The occurrence of a pause of an unusual length heightens its signification as a temporal interval, however small.\textsuperscript{133} The pause can be conceived of as something that introduces an interruption, or in spatial terms, a distance. This break in continuity can be emotionally experienced by participants as either a threatening gap, a benign space, an empty space, a distance too short or long, or a place to hide (the pause is a very efficient hiding place and, by listening to it attentively, something concealed may emerge\textsuperscript{134}). Hebrew words are often heavily condensed with meanings. The word “gap” (halal), can variously mean space, outer space, emptiness, vacuum, death (a casualty of war), and is also the root of the word “window” (halon): a space to see through. The phenomenal meaning of pauses in the clinical encounter is just as varied: pauses are not necessarily empty spaces (although they could be), nor are they uniform. They can be some-thing that moves and changes dynamically, having color and quality: a pause can be anywhere between meditative and stormy, intimate or withdrawn, alive or dead. Consider the following example of the experience of an altered pause between analyst and a patient, reported by Lachmann and Beebe (1996), which they regard to be a heightened affective moment:

\textsuperscript{133} It has been demonstrated that infants, as listeners, exquisitely discern small changes in pace, timbre and intonation (Malloch and T Favrun, 2009).

\textsuperscript{134} Bion (1957) has described such pauses in an analytic dialogue that alerted him to the fact that the patient, who seemed to be doing analytic work, was in fact only going through the motions.
In the treatment of a highly intellectualized man who was dependent on a continuous verbal flow, a long, shared, joint silence became a heightened moment. At this point in the interaction, the verbal content had ushered in a new intimacy, which was then marked by a pause that was much longer than usual. Both analyst and patient extended the pause to savor this new experience. For this patient, the long moment of pausing was felt to be his first direct emotional communication to the analyst, without words. It came to symbolize the possibility of a new kind of shared intimacy.... After this event, the patient had more access to a calmer state in which his verbal flow contained more pauses. There was more room for both patient and analyst to reflect, absorb, and experience. In turn, the mutual regulation was altered. The analyst was more present, both for the patient and for herself. Both patient and analyst increasingly came to be able to describe this altered interactive process in words.

(Lachmann and Beebe, 1996, p. 5).

In this example, the “long, shared, joint silence” was created by extending the expected switching pause into one longer than usual (it is not known from the example whether it was the patient or the analyst who initiated it, though both encountered and sustained it). This illustrates the eloquence and rhetoric of a pause and its dyadic co-construction.

What really occurs during an unusual pause is that participants are unexpectedly propelled into uncertainty – or in Stern’s terms (1985), into a primordial emergent
intersubjective state. Lacking immediate knowledge as to how to understand the situation or proceed, they have no other option but to experience it more viscerally, becoming more aware of body language, sensory impressions, etc., for example, to their shifts of posture, facial expressions, gazes, or lack thereof, breathing, even odor.  

Experiencing silence with an openness and relinquishing the pressing need to immediately know what it is about was advocated repeatedly by analysts who worked with nonverbal or severely disturbed patients (for example Ogden, 1989; Winnicott, 1954/1984c). Like in a painting by Escher, in the silent situation the dialogue is turned inside out: what was logically the background (silence) comes to the fore, changing the participants’ perspective.

Another aspect of pause is that the intrapersonal style of the pause is different for patient and therapist. Pauses are characteristic of analysts’ general listening stance. The pause signifies, in this case, a withholding of immediate action or response, facilitating a slower tempo that conveys an attitude of reflection and careful consideration. Conversely, increasing the tempo of the switching pause by the therapist may signify the advent of an enactment, in which the analyst is unwittingly drawn to a more immediate action.

As an intersubjective phenomenon, the pause represents an emergent – in other words a new - interpersonal event that calls for a new shared experience. The pause thus becomes an aural space – albeit a very small one that transpires rapidly - where issues like patience, acceptance, being with, joining, openness to the

---

135 Coltart (1993) and Khan (1974) emphasize this aspect in their discussion of the silent patient.

136 This is not necessarily negative. In fact, writers like Stern et al. (1998) and Ringstrom (2001) illuminate the vitality of the analyst’s occasional spontaneous, prereflective response.
unknown, creativity, spontaneity and adventurousness (and their respective opposites, e.g. embarrassment, uncertainty, alienation) are enacted in a performative, pre-reflective manner that is experienced as a narrative, or “lived story”.

**Vocal Persona**

The term “persona” refers to a voice or a character in a dramatic or literary work. Originally, the term stands for the mouthpiece of the mask in ancient Greek theatre, through which the voice of the actor was heard (*per* = through, *sona* = sound; see Moses, 1954, Ihde, 2007). It is significant that this term, eventually coming to signify a person or a character, is derived from the voice. Persona can be understood not only as the protagonists, but also as the organizing consciousness of the narrative. For example, the voice of the invisible narrator can be a persona which is differentiated from the author’s personality. The concept of persona responds to questions like: who is the speaking agent in a poem, a work of fiction, a musical performance, or, for that matter, the psychoanalytical clinical setting? Theories on persona problematize the assumption that voices are identical with their speakers, highlighting the idea that, whether concrete or metaphoric, voices are always mediated. In sociology, persona denotes the role that a person assumes or displays in society - one’s presentation in the public sphere, which is intrinsically related to social structure (Goffman, 1959). Modern cultural discourse (for example Foucault, 1970) has further elucidated how issues of power and authority underlie the social roles individuals assume. The concept of persona is also central in Carl Jung’s theory, relating to what in Winnicott’s terminology would be called a “false self”
(1960/1984a) – an organization of behaviors and attitudes that negotiates between personal meanings and social expectations; a way of relating to the outer world that has an adaptive and sometimes pathologically defensive character. An important understanding that emerges from these conceptualizations is that a persona entails a certain manipulation or construction by the speaker, certain choices made by her vis-à-vis her presentation in public – whether these choices are conscious or not. Hence, the persona is essentially a relational concept that, though addressed from different viewpoints – e.g. aesthetic, social and psychological - and for different purposes, is always about a negotiation between self and others, between subjective meanings and the public domain. In these negotiations, the role of the Other as a determining force in the construction of a persona looms large, though the individual’s particular and unique choice of behaviors is equally decisive (and this is precisely the basis of Bakhtin’s dialogism).

In a very different way, the concept of persona as an “organizing consciousness” has been developed in the musical sphere, initially by the music theorist Edward Cone (1974), who conceptualized a sense of a virtual persona that emerges from the musical work, an idea especially intriguing when applied to non-texted music. Differentiated from the person of the musical composer, the musical persona is an illusory sonic agency, a “will” that is embedded in the musical occurrences (for example, musical phrases that “strive” to ascend or descend, a sense of emotional conflict that develops in the music and its resolution, etc.). The idea of musical persona has been much elaborated in music, cultural and semiotic discourse (a representative list is Auslander, 2006; Cumming, 1997, 2000; Echard, 2005;
By elucidating the intrinsic relation between musical meaning and embodied experience, it provided an explanation, different from the more exclusively formal approaches, of how such meaning is generated. The central question here is, in what way do sonic elements – such as musical phrases or sections, or even the totality of the musical work – evoke, through their continuous concatenation, a sense of human dynamic happening and action that is imagined as representing a virtual agentic entity and interpreted anthropomorphically as a character, or a persona?

A specific focus of interest in this discourse, and one that is particularly relevant to the present research, is the musical persona of the performing musician. In her important semiotic study on sonic subjectivity as manifested in musical performance, Cumming (2000) explores the issue of musical identity and the ways in which the performer's subjectivity is expressed, constructed and articulated in the course of their musical playing. The way a violinist interprets a piece of music – and Cumming places great emphasis on the minute, embodied details of performance - conveys, according to Cumming, signs of the performer's subjectivity. Again, this subjectivity is not a naturalistic entity nor a direct expression of a self that exists apriori and emerges through the playing, but a carefully constructed complex (hence, a persona) that is generated through the playing, revealing, in an embodied way, the player's felt and imagined intentions, aesthetic and stylistic preferences and choices, as these coalesce – always in interaction with interpretive listening - into a sense of a public personality.

It is here, in my opinion, that music therapy theory often errs, by inferring a “natural” authenticity to improvisations produced in therapy, without sufficiently taking into account the social and cultural constructedness of such improvisations.
What is the relation between the virtual, fictional character that evolves through the music and the person of the performer? The main issue here concerns the relation between the perceived persona of the performer and the real person who is performing (this obviously raises questions about the problematic inferred existence of an inner “real” self, but these are too vast to go into now). Auslander (2006) for example, investigates performers’ musical personae by applying the dramatic concept of persona to performing musicians (a necessary move since musicians do not usually portray fictional characters\textsuperscript{138}). Auslander points out the gap that exists between an emerging musical virtual persona and the actual person of the performer – indicating, for example, that the musical persona of Bob Dylan does not necessarily reflect the actual person of Bob Dylan. Echard (2005) argues similarly in his study of singer-songwriter Neil Young. Auslander suggests that when we see a musician perform, we are not simply seeing the “real” person playing; as with actors, there is an entity that mediates between the musician and the act of performance. When we hear a musician play, we in fact hear a version of that person constructed for the specific purpose of playing music for an audience under particular circumstances. The musical persona of the performer is thus an entity that is both mediated and stylized by the performer’s value systems, ideals, and choices regarding the expectations and ideals coming from a specific audience as well as the wider cultural context. Musical performance may be thus defined as a person’s representation of self within a certain public musical domain (for example, a concert). Auslander posits that “in musical performance, this representation of self is the direct object of the verb to perform. What musicians perform first and

\textsuperscript{138} Exceptions are of course, operatic personae as well as Lieder’s protagonists.
foremost is not music, but their own identities as musicians, their musical personae” (ibid. p. 102). Now, how does the idea of musical persona contribute to the present psychoanalytic concern?

Vocal Personal in the Psychoanalytic Arena

The association between the psychoanalytic clinical scene and a dramatic or fictional one is quite conceivable and has been often used in the literature. Comparisons of the analytic arena to a theatrical stage (Freud, quoted in Auslander, 2006; Ringstrom, 2001) or a game (Winnicott, 1971), with their attending roles and characters are well known. The question “who is speaking to us” in artistic works and events is formulated in similar terms, though for different purposes, by several analytic authors (see for example Bollas, 1996; Heimann, 1956; Spence, 1982). They envision the analytic conversation as a scenario in which patient and analyst enact different characters through the transference relations. Spence (1982), for example, states:

Only rarely do we find a patient speaking to the analyst in a two-person dialogue; rather, the patient is speaking with multiple voices to a variety of loosely defined 'others', and the conditions of the conversation—

who is speaking to whom and for what reason—become critical determinants of what is being said. The parties, always shifting, are rarely labeled; the analyst (and eventually the patient) must deduce, from the form and content of the utterance, who is being addressed, and he must make this deduction

139 This understanding underlies the work of clinical music therapy, especially with nonverbal patients whose musical improvisation, read as “musical persona”, reflects on their character and self-organization in the interpersonal sphere.
quickly, before a new 'speaker' appears and all clues have vanished. The identity of the 'speaker' and his demands always seem to take precedence over what is being said. Whereas in the usual conversation we take the speakers for granted and focus on the content, in the analytic situation we are often more concerned to know 'who' is 'speaking' and for what reason. (Spence, 1982, pp. 123-124, italics added).

Bollas (ibid.), expands on this notion and suggests that the analyst too, drawn by the patient’s transference into an unconscious scenario, plays different characters, and that the question, initially formulated by Paula Heimann (1956) about the patient's transference— “who is speaking to whom and why now”— can be equally relevant to the analyst: “Who is this speaker, to whom is he speaking, and why?”.

Conceiving of the analytic encounter as a theatrical stage makes it possible to consider the participants as characters (personae) who emerge and populate the analytic drama. These characterological configurations are here-and-now performances of historical scripts -- self-states that were formed in the history of the individual in relation to different significant others. Because of the suggestive frame of the analytic setting, they are enacted in the present in relation to the analytic partner. As Spence succinctly describes it, these “characters” unfold in a vague manner, unlabeled and prone to vanish, clearing the floor for new emerging characters. Now, the only possible way they may become known in the shared interpersonal domain is by means of performance: through behavior, speech,

140 Cumming (2000), too, remarks about the concept of musical personae that the answer to the question “who is speaking to us” is not clearly delineated, nor necessarily vital. More important here is the possibility of asking the question, that facilitates the use of the persona as an interpretant in understanding the dynamics of musical fields.
gesture, posture, and of course, nonlinguistic vocal expression. All these differing modalities comprise the global effect of a persona. Though each modality contributes to the persona its own specificity, phenomenally, they are experienced as inextricable. Still, it is possible, for analytic purposes, to examine them separately.\textsuperscript{141} It is here that the concept of vocal persona makes its entrance.

In the course of dialogue, recurring nonlinguistic vocal happenings can be perceived as organized and stable at a global level, signifying a speaker’s general bearing in terms of affective tone, behavioral patterns, relational attitudes, presentational choices, etc. This global vocal form is a “vocal persona”: a complex synthesis, composed of recurring lower-level phenomena like typical gestures, timbre and rhythmic patterns, that coalesce to imply a singular vocal “character”, a psychological agent or personality that materializes sonically. In ordinary life, we naturally tend to perceive the vocal persona as expressing the speaker, in other words we relate the emerging vocal character to the speaker’s self. But, if we entertain for a moment the ideas coming from persona theory, we can see that the vocal persona is a creation that is not necessarily coextensive with the speaker’s - any speaker’s - total personality. The usefulness of the concept of vocal persona for the psychoanalytic endeavor is suggested in the following:

1. As a mediated phenomenon, the vocal persona can be apprehended as a sonic \textit{icon}\textsuperscript{142}, i.e. a relatively stylized sonic structure that, more than revealing the

\textsuperscript{141} The media and the arts have provided fascinating scenarios in which personae appear in exclusively aural ways, for example, as fictional characters in radio dramas, or exclusively visual - through pantomime or in the silent cinema. These instances prove that characters can be portrayed through single modalities.

\textsuperscript{142} As we have seen in chapter II, sonic icons are related to signifieds by way of resemblance, meaning that they are farther removed from them than indices, which are immediate and contiguous in
The speaker’s “real” self or character, attests to the speaker’s subjectivity by revealing her intentions, her negotiations between inner self and outside world, her preferences and choices of articulation of subjective states in the public sphere. The vocal persona can be thought of as the external outcome of inner self-regulation processes (Beebe and Lachmann, 2003). The individual mode of regulating inner occurrences, like arousal, affect, expectancies, fantasy, perception of inner states and of how the individual is affected by her own behavior, have outward signs: these are manifested in an embodied vocal performance. For example, when a speaker manifests a decidedly subdued vocal persona, it is through certain vocal choices that this is displayed, perhaps through a recurring muffled vocal timbre, monotonous intonation and vocal gestures, a narrow vocal range, avoidance of vocal stress etc. This vocal icon of submissiveness can be heard not just as an expression of a passive subjective state, but as a deliberate (though unconscious) construction, made by the speaker in relation to significant others in the course of her life history, that is performed in an embodied vocal manner – through concrete audible vocal acts.\footnote{It is emphasized however, that I am not suggesting here a formal, external analysis of vocal signs. As I have maintained through a phenomenological approach is that the understanding and interpretations of these occurs through prereflective identificatory processes.}

Bakhtin’s concept of self-authoring helps to better understand this. As we have seen, the individual’s dialogue with existence is seen by Bakhtin as a continual process of self-articulation, or, in Holquist’s (1990) phrasing; selfhood is understood as a form of authorship. The vocal persona is thus the vocal manifestation of the authorial self. Musical improvisers often think of themselves as composers-in-the-moment, and in a certain way, the vocal persona can be seen as an amalgamation of relation to their signifieds. For example, a picture of a smiling person or a road sign (icons) are more deferred and often more stylized than a person’s smile (index).
author, composer, narrator and musical performer (this echoes Merleau-Ponty’s poetic view of speech as “singing the world’s praises”. See chapter II, p.36).

Consider the following example from an analysis (Sandler, 1976), at the point in which the analyst becomes aware that, in relation to a certain patient, he is doing much more talking than usual:

I noticed at once the urge to talk during the session and became aware that the patient, by a slight inflexion of his voice, succeeded in ending every sentence with an interrogation, although he did not usually formulate a direct question. This gave me the opportunity to point out to him what he was doing (he was quite unaware of it, just as I had been unaware of it in him) and to show him how much he needed to have me reassure him by talking. He then recalled how he would feel extremely anxious as a child when his father returned home from work, and would compulsively engage his father in conversation, asking him many questions in order to be reassured that his father was not angry with him. His father had been a professional fighter, was very violent, and the patient was terrified of him but needed his father's admiration and love, to be the preferred child [...] He told me that his father had the habit of not listening and not responding, and how frightening this was. The patient then realized that from early childhood onwards he had developed the trick of asking questions without directly asking them, and this had become part of his character, being intensified in situations where he feared
disapproval and needed supplies of reassurance from authority figures. (Sandler, 1976, pp. 44-45, italics are mine).

This clinical illustration exemplifies the idea of vocal “self-authoring”, through the way that the patient manipulated his voice in a defensive manner, by stylizing his speech with interrogating vocal gestures (probably inflected with ascending pitch); and through making the music of interrogation part of his vocal character.

2. The idea of vocal persona illuminates the sonic aspect of psychological defense mechanisms such as dissociation, splitting, projective identification and isolation. The partial autonomy of the nonlinguistic vocal dimension facilitates its (unconscious) use in the expression of repressed, denied or unaccepted aspects or feelings of the speaker’s self. Thus, the voice can become a safe place for the expression of emotions that are perceived as dangerous or forbidden. Of course, these mechanisms are unconscious. In the process of becoming acquainted to the patient, a close listening to her vocal persona can reveal how, at times, these sonic aspects are characterologically organized in the voice.

Vocal persona also effectively relates to theories that assume the self to be a complex of multiple self-states often experienced in fragmented manner, rather than a unitary entity (Bromberg, 1994). According to these conceptions, individuals experience themselves through different selves or self-states that appear and disappear according to circumstance. Bromberg suggests that each self-state has “its own narrative, its own memory configuration, its own perceptual reality, and its

---

*Stern (1985) has argued that the nonverbal dimension is the private, deniable version of the self, that can thus be used to express socially unaccepted feelings and themes (the verbal version being the “official” one).
own style of relatedness to others. It is not simply something one feels—it is who one is, at least at certain times…” (p. 520). Bromberg believes that analytic exploration must enter in dialogue with the patient’s multiplicity of different self-narratives and perceptual realities “… each of which speaks with its own voice. To put it a bit more poetically, an analyst must be constantly negotiating with a range of self-states with different voices” (Bromberg, 1994; pp. 521-522, italics added). The concept of vocal persona helps us understand that these “voices” are not just a poetic metaphor, but rather that they have concrete sonic manifestations that come through to the analytic listener in the form of different vocal personae.

3. Persona theory, as well as Bakhtin’s idea of addressivity, imply that personae and audiences are fundamentally wedded: personae are always articulated in relation to readers, viewers or listeners. And just as audiences infer a character from the performer’s persona, so do speakers address themselves to virtual listeners in the process of generating their personae. This is perhaps one of the most important aspects of vocal persona in the analytic context: possibly more important than asking “who is it that is speaking?” is asking “who is it that is listening?”. In other words, who is the implied listener in the speaker’s utterance (to whom is the speaker addressing herself at this moment) and how is this listener affecting the organization of the speaker's vocal persona?”. I have maintained above that all vocal phenomena are understood as intersecting past and present actors and scenarios. The vocal personae emerging in psychoanalysis amalgamate, in their address, two audiences: the one is historical (introjected listening others), the other is the present analytic listener. In the example given by Sandler above, we see how the theme of authority
- in the past, through the father and, in the present, in relation to the analyst - is elaborated by the speaker through his vocal persona. Similarly, vocal personae manifesting competitiveness, provocativeness, cooperation, submission, hesitancy, anxiety, etc., reflect vocal characterological organizations addressing virtual and concrete listeners. Other major relational themes reflected in vocal personae are attitudes towards intimacy, dependence and degrees of permeability in relation to others.

4. Finally, some consideration must be given to the analyst’s vocal persona. The analyst's role has been a central theme of analytic attention. Various related issues have been debated, like the authority of the analyst, her responsiveness and countertransference – but, to my knowledge, almost no attention has been paid to the manifestations of these themes in the vocal plane (though some attention has been given to the performance of the analyst through the relatively new concept of enactment). It should not come as a surprise that the analyst's general attitudes and implicit value systems towards her role are reflected vocally. Thus, the vocal persona of the analyst, composed of her attitude and value system regarding her professional role, his social and personal value systems as well as her own internalized dialogical patterns, affect her vocal persona.145

---

145 This is exemplified, albeit in a highly stereotyped manner, in cinematic portrayals of therapists. Compare, for example, the vocal persona of the analyst as embodied by Ingrid Bergman in Hitchcock’s “Spellbound”, versus Robin William’s rendition of the therapist’s role in “Good Will Hunting”. Placing these films’ implied attitudes about therapy aside (a fascinating theme in itself), it is interesting to study how these professional roles are constructed vocally.
**Vocal Duet**

Until now we have considered vocal phenomena occurring at the level of the individual. As I have maintained, these are fundamentally dyadic. One aspect of this dyadic nature is manifested in the implicit and inferred manner in which the virtual Other permeates the individuals’ utterance. The concept of vocal duet which I now address is dyadic in an explicit and material way. We are here considering the vocal happenings that transpire between the two analytic participants. I have called this dyadic phenomenon vocal duet in order to highlight the mutuality and interpenetration of the partners in creating textures, forms and patterns that echo a two-person musical interaction. The “music” of the vocal duet exemplifies a mode of dyadic participation that generates meaning in an embodied and prereflective manner.

The affinity between the analytic relation and a musical improvisational duet is not foreign to analytic conceptualization, especially in recent decades (for example Benjamin, 2002; Knoblauch, 2000; Lachmann, 2001, 2004; Urban, 1993) and the similarities between the two are often referred to in infant research (Gratier and Apter-Danon, 2009). Music theory addresses the music of speech and conversation (Lidov, 2005; Rousseau), and equally, musical improvisation as an act of “saying something” (Monson, 1995). These metaphors indicate the organized and semantic nature of sonic dyadic exchanges in both music and conversation.
Vocal duet constitutes a level of communication that, though occurring simultaneously with verbal communication, has different principles of organization and operation and is perceived and understood mostly out of awareness. As discussed above, Knoblauch (1997) refers to this flow of vocal behavior as a continuous-process, in contrast with discrete communication. He explains that with a continuous-process, the rate of communication is rapid and, at any point in time, bidirectional. Unlike verbal exchanges, both participants are here continuously active and each has the opportunity to modify their own actions immediately and without waiting for the partner to complete a turn or produce a discrete signal. Knoblauch indicates that, in this situation, it is difficult to say who initiates a communication or who responds to whom, as each participant shapes the experience and activity of the other. As explained by Beebe and Lachmann (2003), in a bidirectional system each person’s behavior is predictable from (contingent to), not “caused by” that of the other. Both participants influence and are being influenced by the partner’s words and actions. In addition to mutual regulation, Beebe and Lachmann emphasize processes of self-regulation that include self-predictability, regulation of arousal, previously established expectancies, symbolic elaboration, fantasy and projection and the perception and articulation of inner states. Each person both coordinates with the partner and, at the same time, regulates their inner state. Self- and interactive regulation are concurrent and reciprocal processes, each affecting the unfolding of the other. As we have seen
above, in the section on vocal persona, processes of self regulation appear on the surface and are manifest in the vocal productions of speakers.

The resulting dyadic vocal texture, or duet, reflects the nature and form of the partners’ reciprocity. This process is thus strikingly similar to what happens in an improvised musical duet (it is important to emphasize that reciprocity does not imply symmetry: the vocal duet does not imply an identical form or style of participation).

*Principles of Operation*

The vocal duet operates largely through the principles described in section 5.3. on this chapter. Motivated by a need for correspondence and coordination, mechanisms of repetition (including imitation, matching, synchronization) and change are central. At the nonverbal level, the interlocutors participate in a moment-by-moment synchronization of their continuous rhythmic-vocal and auditory behavior, in their vocal inflection and timbre, pause and silence, tempo and vocal gesture. In this manner, partners prereflectively establish expectancies and norms, that in turn define boundaries and constraints, create stable patterns and forms of behavior as well as ranges of freedom for new initiatives. These processes were observed in mother-infant interaction, compared to processes of music improvisation and described as an intricately arranged duet for voice, eyes, face and hands, “as tightly organized as the performance of well-matched and highly practical dancers or musicians” (Trevarthen, in Urban, 1993, p.239). Microanalysis of recordings of proto-conversations, as they are called, reveal great similarities between infant and adult rhythms and expressivity. For instance, mother and infant are drawn to join in with one another on a shared rhythmic beat. Urban elucidate
that when mother and baby attempt to synchronize their vocalizations, the duration of the pause designed to catch up with the other and establish the beat, is the same as that of various pre-beat signals in adult conversation, the article before a noun or the up-beat of a conductor's baton, etc. (ibid.). Analyses of musical improvisation show the same traits: a fluid, bidirectional, continuous, self- and other influence and response, motivated by the partners' constant search for synchronization, a common beat, “feel” or “groove” (Berliner, 1994; Monson, 1996).

**Participation and Roles**

The bi-directionality of communication that Beebe and Lachmann describe has implications regarding how communication in analysis is conceived. In the verbal mode, discrete communication focuses on sequential turn taking – while one participant talks, the other listens, then they exchange roles. The analyst's role is conceived predominantly as an external observer who listens more than he talks. The present bidirectional model, focusing on co-action and the continuous, simultaneous nature of communication, views the analyst rather as a participant, who, in an out of awareness mode, is continuously affecting the patient as well continually affected by the patient. In the relational and intersubjective psychoanalytic movements there is an increasing understanding of the analyst as participant-observer. This point will be addressed further in chapter IV.

Beebe and Lachmann (1996, 2002) examined the process and patterning of dyadic interaction and identified salient forms of exchange and regulation:

---

146 Visual contact in face-to-face communication also has an important role in ordinary and musical communication. However, the nature of sound permits communication without visual contact, as in the psychoanalytic setting.
a. Ongoing regulations - these are the characteristic, expectable patterns of repeated interactions in the treatment situation, manifested in a wide range of nonverbal behaviors, such as postural and facial interchanges, intonations and tone of voice, and greeting and parting rituals, as well as through verbal exchanges. In the present context, my concern is with the vocal aspect of ongoing regulations.

b. Disruption and repair – (discussed more extensively in the section on pause). This is a specific instance of the principle of ongoing regulations. It organizes violations of expectancies and ensuing efforts to resolve such breaches. Changes in vocal behaviors effect a disruption of expectancies and present the dyad with a problem that needs to be resolved. The relational pattern of “chase and dodge” was identified by Beebe and Lachmann (1988) as one example of a stabilized disruptive pattern, where one partner, typically the mother, continually overstimulates the baby who develops in response typical withdrawal (dodging) behaviors. It is easy to imagine this pattern of interaction in adult relationships.

c. Heightened affective moments. This principle refers to interactions that are emotionally experienced as unique, powerful states of transformation, either positive or negative (these are Stern’s “now moments” [1998]). These authors see such heightened affective moments in patient-analyst interaction as providing opportunities for new experiences. These moments are often manifested along nonverbal parameters.

The patterns that Beebe and Lachmann describe can be paralleled to musical forms such as unison, fugue, counterpoint, heterophony, polyphony, theme and variations, call and response, and so on. Conceived as a dyadic creation of quasi-musical forms,
the vocal duet reflects vocal patterns of corresponding and relating that, as suggested, reflect a blend of both past and present interactive patterns. Attending to these, as they are manifested in the typical vocal roles that participants assume, can enrich and deepen the analytic investigation of relational meanings as enacted in the present. For example, in a “chase and dodge” vocal form – let us call it a musical fugue – who is the “chaser” and who, typically, the “dodger”? Analysts often report the experience of being “dodged” by patients (but interestingly, according to popular lore, analysts are almost always perceived as “dodging” patients' direct questions). These forms are manifested vocally through variations in pause, matching and mismatching of vocal timbre and gesture, and so forth.

An additional point is that, as I have indicated above, modes of participation are not identical. The analyst's verbal output is characteristically smaller than that of the patient and her pace is generally slower, representing a stance of deliberate, intentional listening and reflectivity. A musical analogy that comes to mind is the form of solo and accompaniment. Conceiving of the analyst as an accompanist and of the patient as a soloist can reveal another facet of the form of interaction, and further illuminate important relational aspects. As we know, there are many forms of musical accompaniment – it can be loose and minimalistic, or dense and tight, receding in the background or asserting more presence. It is informative, in this respect, to listen to different performances of the same song or musical work. Differently arranged accompaniments often succeed in producing totally different dialogical textures between soloist and accompanist, thus generating strikingly

\[147\]  This metaphor was also used by Rose (2004).
different meanings for the same song. Ordinarily, the accompaniment maintains the harmonic structure and the rhythmic beat in a steady manner, allowing the soloist a relative freedom of expression, releasing her from structural constraints. This is a good analogue for the analyst's steadiness in holding the clinical setting, endowing the patient not only with a freedom of expression (like free association), but with a psychological experience of “being held”. 148 Balint (1968) referred to the analyst's presence as “an environment that accepts and consents to sustain and carry the patient like the earth or the water sustains and carries a man who entrusts his weight to them. In contrast to ordinary objects, especially to ordinary human objects, no action is expected from these primary objects or substances; yet they must be there and must — tacitly or explicitly — consent to be used, otherwise the patient cannot achieve any change. “... The substance, the analyst, must not resist, must consent, must not give rise to too much friction, must accept and carry the patient for a while ...” (p. 144).

To the metaphor of the analyst as earth or air, we can now add that of the analyst as musical accompanist that “holds” the patient through a steady, predictable and reliable “accompaniment”, manifested in her vocal-rhythmic behavior. Knoblauch (2000), describes his periodic vocal interventions with a silent patient that, more than conveying discursive meaning, signaled to the patient his continued presence and listening. Obviously, there may be other styles of accompaniment. As we know, some musical accompaniments have more presence and individuality (see for example Monson's [1996] study of different styles of the rhythm section in jazz improvisations). Similarly, the analyst-qua-accompanist sometimes displays a more

148 Tustin (1986) has aptly called the analytic holding “the rhythm of safety”.

218
interventive, or more participatory style, whether because he senses that a closer accompaniment is needed by the patient (Knoblauch [1997], for example describes his nonverbal vocal interventions with patients as means of synchronizing, calming, or “picking up” distressed patients), or because he was drawn unawares by the patient into a more active participation (see the example of Sandler above). Observing the form of accompaniment and the changes it undergoes can facilitate the understanding of important themes in the transference. The vocal responsiveness of the analyst-accompanist to the different needs of the patient underlies the cultivation of trust, interdependence and creative possibility in the interacting dyad. Infant researchers have demonstrated that the responsiveness of the mother, manifested in her vitality, and her ability to respond with flexibility and variation, inspires feelings of security, belonging, trust and creativity. Similarly, musicians talk about these vital feelings as occurring through successful improvisations.

4. I have proposed in above that one of the organizing principles of the nonlinguistic dimension is its narrative ability. Trevarthen and Gratier (2008), investigating the narrative value of proto-conversations between mothers and infants, suggest to hear these first narratives as nonverbal musical stories that mothers and babies spin together. These narratives are about “finding oneself in the other”, about “the special ways in which understanding the other means losing oneself or about the

149 I have elsewhere (1996) demonstrated that, in songs created by children in music therapy, the musical use of the therapist in the process of creating the song reflected specific object-relation modes. These were manifest in the changing role of the therapist in relation to the situation. Generally, it was found that in characteristic paranoid-schizoid transferences, the therapist’s musical presence in processes of composing and performing with the patient was greater. Where the transference was more secure (depressive stage), the child was better able to “play alone” in the presence of the therapist (Winnicott, 1958/1984) whose musical interventions and participation were more removed and in the background.
reciprocal attribution of intentionality and agency”. The felt immediacy of the musical forms that emerge in the dialogue “draws the infant into a fictionalized space where known shapes of time are recounted and conjure up emotions of belonging. Meaning grows out of the experienced flow of shared embodied narratives” (ibid. p. 138). These poetic remarks about wordless narratives hold equally true in the adult analytic exchange. As I have already pointed out, I propose that the sense of aboutness in the vocal duet is self-referential: it is a vocal-musical way of telling each other about “what is happening now between us”. It is a musical story about how it feels to be together, how we co-create our togetherness, and what do we do in order to be-with one another. The referent of the vocal duet is the relation between the speakers. These processes are deeply linked to issues of agency, authority and influence, belonging and self esteem, the degree of permeability to others, the ability to rely and depend on others and be conversely dependable.

**Vocal Ambience**

The most complex – and perhaps most intangible - vocal phenomenon I suggest is Vocal ambience. It is a wide-lens aural impression denoting the texture of the dyadic interchange as-a-whole. Although many analytic writers referred to the ambience of the analytic encounter, few have articulated its nonlinguistic vocal aspect or its musical sense (but see Knoblauch, 2000; Lachmann, 2001; Rose, 2004; Stein, 2007). Perceiving the nature of vocal ambience requires an aural zooming-out regarding the particulars of the exchange and the development of a sweeping attention to the totality of the vocal field. (A similar experience takes place when, while looking at a
painting, one steps back to get a wider, more global perspective on the work. The wide lens perspective reveals aspects of the work that are impossible to see from a closer perspective.

Vocal ambience is effected by the total flow and interaction of all vocal elements and phenomena coming from both analytic participants, ranging from the most simple to the more complex syntheses of vocal personae and vocal duet. These coalesce in perception into what can be thought of as a “music”, a sonorous place, space or environment with a characteristic texture and tone. It is this blend of the various timbres, gestures and rhythms, ranges of pitch, melodic contour and variation that creates the vocal ambience. More than engendering a semantic value, ambience characteristically creates a sonic mood that hangs, as it were, in the air. Although phenomenally experienced as “state”, it is in fact a dynamic phenomenon (though somewhat more stable than the dynamics of the vocal duet), sensitively reflecting the emotional texture of the unfolding dialogue. Ambience is constantly in the making, created anew in the course of each analytic encounter. Curiously, vocal ambience works in two opposed directions. On one hand, it is generated by the analytic participants, but, becoming an autonomous entity it affects the participants, since it is phenomenally experienced as a medium in which they are enveloped or immersed - though mostly out of awareness.150 Thus, Vocal ambience works much like an eco-system: the reciprocal relationship between the different parts of the

150 We have seen above that the voice has been conceived psychoanalytically as a sonorous medium akin to a sonorous bath, or an audio-phonic envelope (Anzieu 1979, 1990 and Lecourt, 1990). While Anzieu and Lecourt engage the baby’s experience, I here suggest that in adult treatment, both participants are similarly immersed in a vocal ambience of their own making, jointly produced by their voices and rhythms. Another notion relevant to sound as ambience is Freud’s idea of the regressive “oceanic feeling” (1930), taken by Kohut (1957) to qualify the experience of listening to music. The metaphor of music as an ocean indicates, too, the phenomenal experience of sound as something that surrounds the listener.
system create phenomena that in turn affect the parts. The experience of the sonic
as a medium is twofold and circular: the partners generate an ambience, and, while
being immersed in it (and experiencing its phenomenal givennes), they are affected
by and respond to it, thus continually shaping and being shaped by it.

The ambience is obviously not composed solely of vocal elements\(^{151}\), but I suggest
that nonlinguistic vocal-temporal parameters play a central role here, especially
since bodily motion and facial expression, normally contributing to the general
ambience, are greatly reduced in the (classical) analytic situation. Thus, the general
emotional tone of the situation is reflected by the gradual, imperceptible, continual
shifts of the vocal parameters, perceived globally, atmospherically and out of
awareness. The perception of vocal ambience as supra-entity does not preclude the
simultaneous perception of other, lower-level sonic occurrences.\(^{152}\)

Vocal ambience is a different category than vocal duet. Although both are
generated by dyadic interaction, the phenomenon of vocal duet is more local,
focusing on the level of dyadic \textit{exchange}. In the vocal duet, attention is paid to the
continuous to-and-fro between partners and to the simultaneity of response: it is
about the moment-to-moment negotiation of relational intention and influence that
occur through the voice. In ambience, the emphasis is not on the exchange but on
the \textit{global effect}, that seems to have acquired a certain independence from local

\(^{151}\) In this context, it is difficult to ignore the influence of the physical setting, including the therapist’s
room, the time of day, the season, weather, and so forth, on ambience. Although these contributing
factors are always present, it is still possible to speak of the unique ambience of a specific analytic
meeting, as something created anew in each encounter by the analytic partners.

\(^{152}\) In chapter II I have discussed the difference between focused attention and global attention. As we
have seen, attention ordinarily oscillates between these two modes, but can also hold both aural
perspectives at once.
vocal behaviors. Vocal ambience is not directed at anyone, but is rather a sonic flow that seems to have a life of its own.\footnote{Ambience is a central phenomena in group dynamics and, though it is often impossible to tell who is its source nor how it developed, members of the group feel that they are “in” an ambience or atmosphere and are influenced by it. This is explained by the phenomenon of social contagion that leads to a diffuse, collective experience created through the group unconscious.}

**Vocal Ambience and the Intersubjective Analytic Third**

Ogden’s concept of the “intersubjective analytic third” (1995) is useful in understanding vocal ambience in an analytic context. He defines it as follows:

The intersubjective analytic third is understood as a third subject created by the unconscious interplay of analyst and analysand; at the same time, the analyst and analysand qua analyst and analysand are generated in the act of creating the analytic third (there is no analyst, no analysand, no analysis, aside from the process through which the analytic third is generated). The new subjectivity (the analytic third) stands in dialectical tension with the individual subjectivities of analyst and analysand. The intersubjective analytic third is not conceived of as a static entity; rather, it is understood as an evolving experience that is continually in a state of flux as the intersubjectivity of the analytic process is transformed by the understandings generated by the analytic pair. (Ogden, 1995, p. 696, Italics added).

I take the intersubjective third to be not some mysterious esoteric entity, but the interactional dyadic fabric as it is manifested through subtle embodied behaviors and responses. From this point of view, vocal ambience reflects, vocally and aurally, the intersubjective third. Ogden is here offering two relevant ideas. First, the
dialectical interplay of two individual entities, that create a *unified third entity*, which in turn, affects the individual subjectivity of the participants. Although I see it as dialogical rather than dialectical (there is no neat dialectic closure in the relation, but rather a messy unfinalizability to the interplay, as argued by Bakhtin), this conceptualization facilitates our ability to *hear* the third entity produced by the analytic interaction.

Second, Ogden emphasizes that the third entity is an evolving experience, i.e., an occurrence that (in my terms), happens pre-reflectively and unconsciously in the moment, and is continuous, motional, and dynamic. The intersubjective experience in the moment-to-moment of the analytic encounter is a theme that Ogden continuously struggles to express in his writings (see for example 1995, 1997, 1999a, 1999b, 2001). It is not accidental that he refers to it as “the music of what happens” (1999a). The use of this metaphor implies the global, dynamic, polyphonic aspect of the vocal ambience and the affinities of the general structure of analytic experience with the structure of musical experience. This theme will be further address in the next chapter.

*Vocal Ambience and Mood*

Bollas is another psychoanalytic author who contributes to the understanding of vocal ambience. In addressing the experience of the analytic encounter, he elucidates the concept of mood. Although Bollas focuses on the mood of the individual patient (and not, as I suggest here, on a shared dyadic mood), he conceptualizes it mostly as the outcome of a dyadic experience. In the present
study, my concept of vocal ambience places greater emphasis on the mutual, though asymmetric, contribution of patient and analyst to the joint analytic mood.

Bollas addresses the nature and function of moods in the patient, suggesting that analysands establish crucial states-of-being in the transference, primarily through the living through of a mood experience (1984). Here too, the emphasis is on an affective state that is lived-through; an existential sense of being which is made present in the clinical setting. Bollas’ conceptualizations add three important points:

a. Moods are affective states that reflect an existential memory of being-with, a non-cognitive memory of a meaningful rapport with an Other (the mother). Bollas talks about a sense of existential being that is born through being-with – the first relationship as a transformative relation. In this experience, the subject’s sense of existence is historically constituted by the other’s (the parental figure’s) mode of being-with her. Being-with is the ongoing communication between infant and mother, where the mother’s task is to provide the infant with an experience of continuity of being. Her handling and the infant’s state of being are prior to the infant’s ability to process his existence through cognitive or symbolic means. Living through moods in the analytic encounter is thus an unconscious reconstruction of a historic, dyadic mode of being-with, unconsciously drawing the analyst to play a historical role in a mood – or, as I here call it, ambience - that, shared and lived together in the clinical setting, informs the analyst of an existential quality that belongs to the patient. Analytic listening and reflecting on the nature and feel of this co-constructed ambience facilitates access to important information that is not
available to the patient in declarative or episodic memory and cannot be thought, let alone verbalized.

b. A second idea that emerges in Bollas' conceptualization is that moods, as mnemonic traces of dyadic happenings, suggest a *location* or a space, an internalized sense of environment. Individuals are “in” moods, and their interlocutors perceive their being not completely present, but partially “somewhere else”. The notion of mood as place is important because, as we shall see, vocal ambience is a co-construction of a sonorous space, a human vocal environment, an eco-system (or perhaps, an *echo*-system) characterized with boundaries, specific textures and profound existential meanings (reminiscent of the Heideggerian Dasein).

c. Bollas significantly conceptualizes the ongoing, sentient experience of being-with as an *aesthetics of being*, referring to the specific maternal manner of handling the baby as a “maternal aesthetics”. He suggests that

the mother’s idiom of care and infant’s experience of this handling is the first human aesthetic. It is the most profound occasion where the content of the self is formed and transformed by the environment. The uncanny pleasure of being held by a poem, a composition, a painting, or, for that matter, any object, rests on those moments ... when the infant's internal world is given form by the mother. This first human aesthetic informs the development of personal character (the utterance of self through the manner of being rather than the representations of the mind) and will predispose all future aesthetic experiences that place the person in subjective rapport with an object.  

From this point of view, the analyst contributes to the analytic encounter, and specifically to the vocal ambience, her specific “analytic aesthetics”, in her way of handling and participating in the analysand's mood.

*Ambience in Music – Sonorous Eco-Systems and Grooves*

Bollas' conceptualization of mood as reflecting an aesthetics of being resonates with vocal ambience as a music and, indeed, ambience is a phenomenon that emerges in all musical objects. It is purposefully created in musical compositions, and emerges spontaneously in musical improvisations. From a compositional point of view, ambience is a constructed textural form, born out of the continuous stream of patterning and interaction among smaller elements and simpler structures. Lower level processes bring forth higher level processes that give rise to an overall sonority, a continuing sound texture rising from the musical work as a whole (Di Scipio, 2002). The musical genres of “soundscape” and “ambient music” (respectively coined by the composer and environmentalist R. Murray Schaffer [1993], and the musician and composer Brian Eno), that evolved in the last decades in contemporary and electronic music, and which were adopted by ethnomusicology and other phenomenological approaches to music are here relevant. While ambience or atmosphere can be ascribed to a work of music of any style, in the last decades a special focus has been placed on this unique musical effect. Musical ambience or atmosphere, brought to aesthetic heights for example by impressionist composers like Debussy and Ravel, evolved in modern electronic music to become a primary theme in musical works, as opposed to the more traditional primacy of discrete and structural elements like pitch, melody and harmony. The musical works produced by
these composers create sonic vistas, impressions and moods, or, as their generic terms have coined them, soundscapes and sonic ambiances. Here, again, the connection between sonic ambience and ecology emerges. As Di Scipio (ibid.) explains, through the production of complex sound textures and emergent sonorities, these works reveal a radical approach to music as a reflection on man and culture. This echoes Ballas’ idea of mood as reflecting early environmental experiences. From a systems-theory perspective, sonic objects, whether musical or conversational, can extend into issues of ecology through the notion of the sonic field working as a kind of ecosystem, namely, as reflecting the continual sonic exchange between human objects in their immediate surroundings and their own history.

Another musical phenomenon that is relevant to the understanding of vocal ambience is the concept of “groove”, coming from Jazz improvisation. Groove denotes a dynamic co-action of improvisational partners in an ongoing search for a sense of togetherness. But, more importantly, it is an aesthetic concept describing the atmosphere of the improvisation in terms of cohesion, energy and “the something more” that occurs between coordinating participants and exceeds what individual musical behaviors can account for. Conceptualizing jazz improvisation as a conversation (1996), Ingrid Monson defines groove:

A jazz band provides a framework for musical interaction among players who take as their goal the achievement of a groove or feeling – something that unites the improvisational roles of the (musicians) ... into a satisfying musical whole. Jazz musicians discuss common experiences of soul, warmth, and
emotional expressivity that have much to do with the ineffable and
unpredictable musical chemistry among players, who take great delight in
recounting their most quirky and magical moments in performance ...  

Monson reports that Jazz musicians Richard Davis and Kenny Washington
emphasized the interpersonal aspect of groove by comparing it to “walking down
the street” with someone. Davis likened groove to a romantic or familial relationship
while Washington saw it as walking arm-in-arm with someone (p.68). In a way,
groove can be taken to be a musical term that denotes something similar to Bollas’
aesthetic moment: grooving is an aesthetic of being and feeling together, an
existential, temporal experience of searching, communicating and mutually handling
and co-constructing sonic environments.

Vocal Ambience in the Psychoanalytic Arena

How does the idea of vocal ambience contribute to psychoanalytic concerns?

1. As we have seen, many analytic writers construe the analytic voyage as a lived
experience, viewing it as a co-creation of analyst and patient (see especially Balint,
1968; Bollas, 1987a, 1989; Joseph, 1985; Khan, 1974a; Ogden, 1999a, 1999b; Stern,
2004; Winnicott, 1971). This co-living in the analytic environment is paralleled to
lived experiences in the early environment (mother) that the patient has
internalized. In this environment, these experiences are not remembered in a
cognitive symbolic way and can be only enacted by unconsciously mapping them
onto the analytic situation. The analytic situation presents an opportunity for
revealing and understanding these early, untold stories. In these views, the
countertransference represents the unconscious ways in which the analyst is drawn to play a role in this historic scene. On the other hand, this is obviously not an exact replica of the historic situation.

Thus vocal ambience, the sonorous manifestation of the global analytic situation as lived experience, reflects the early infant-mother eco-system carried over to a here and now situation. This is exceptionally important with nonverbal patients who cannot verbally report their inner experience, or with verbal patients where this early, archaic experience was not symbolized. The shared living-through of this experience is a prerequisite for its consequent symbolization.

2. Attending to the vocal ambience means asking questions about the general sound, or texture, of the conversation. Focusing on the manner in which vocal elements coalesce into a global texture facilitates asking what kind of “music” is unfolding in this particular encounter? Or, what kind of “groove” is going on in our conversation? These questions can guide the participants’ attention to elusive meanings that reside in a covert manner in the vocal level but are not manifestly discussed.

Characterizing the vocal ambience can qualify the form of togetherness, the particular aesthetic of being-with between patient and therapist at a given encounter (or, when recurring, as a stable style). In other words, the quality of vocal ambience is an index of the form and quality of various relational themes: the shape and quality of being-with, the mutual, co-constructed sense of belonging, issues of trust and mistrust, varying attitudes to authority, tradition and innovation. At times, the aesthetics of vocal co-being is one of harmony, manifested, for example, by a
close matching of rhythm, timbre, intonation, etc. Although this is usually an index of a positive sense of co-being, sometimes an overly harmonious vocal ambience can signify the patient's – or analyst's - avoidance of confrontation, conflict or aggression (this is often the case in group therapy). At other times, the vocal ambience is wrought with tension and dissonance. These could possibly be manifested sonically through tense vocal timbres, jagged gestures, a sense of vocal “pulling and pushing”, sudden vocal irruptions and bursts, mismatching and poor synchronization. Of course, tension and dissonance can be manifested in manifold other manners, including silence. In addition, vocal ambience accurately reflects the vitality of the encounter: The sense of the aliveness or deadness of the analytic encounter is an index that some authors are keenly attentive to (for example Ogden, 1995), and which is, again, manifested mainly through nonverbal parameters.

3. The autonomous texture of vocal ambience (since it is, unlike words, a continuous, unsegmented flow) and its aural presentation as a separate entity, makes its relationship to the verbal plane varied and complex. Vocal ambience sometimes seems to support verbal contents and even reinforce their meanings. At these times, it is hardly noticeable, as is usually the case when there is a coherence between form and content in speech. At other times, vocal ambience may express different, even contradicting, currents than those conveyed verbally. The field of music has vastly exploited and researched this phenomenon as an artistic tool and a powerful rhetorical device in opera, lieder and film-music (Cone, 1974; Epstein, 1993; Kramer, 2002; Langer, 1970; Monelle, 2000; Storr, 1992). By expressing certain dynamics independently from verbal or visual happenings, the sonic sphere
creates a space or gap between the various levels of signification, enriching the emotional and psychological power of the event. Similarly, paying conscious attention to the nature of the vocal ambience (as the “soundtrack” or, sonic background), and to its relation to the verbal and behavioral levels opens new and important spaces of analytic exploration.

* * *

As argued by phenomenology, phenomena can only be understood in light of the structure of their perceptual counterparts. Having addressed the various vocal principles and phenomena in this proposed model of the psychoanalytic voice, it is time to elucidate sound's corresponding perceptual mechanism – listening.
Chapter IV
Listening

IV.1 Introduction

Although appearing as a distinct topic, the subject of listening is far from being new to the present study: as we have seen, the phenomenological inquiry of a phenomenon is at the same time an investigation of the structure of its perception, hence, core issues related to listening have been addressed throughout this research. In chapter II, I have endeavored, through a phenomenological epoché, to make the nonlinguistic voice audible. By decentering logos and attending to the sounds that exceed the verbal voice, the structural and phenomenal nature of the fluid auditory realm was discussed. Subsequently, music aesthetics and semiotics were engaged to explicate the ways in which nonverbal sonic happenings are organized and prereflectively perceived (heard) as a motional, indexical and iconic semiotic system – the specific organizing principles and forms of which were described in chapter III. Finally, I turned to dialogism as the overarching theoretical framework that directs listening to the polyphonic soundscape – the heteroglossia – that constitutes utterance in human interaction.

Similarly, the different psychoanalytic interpretations (“hearings”) of the voice that I suggested in the literature survey in Chapter III are as much about listening as they are about the voice. What actually surfaced in the survey are the preconceptions and biases underlying analytic listening and their embeddedness in ever-changing historical, cultural and theoretical universes. Thus, although that chapter explicitly
addressed the materiality of vocal phenomena, the implications that came forth concern listening. The central understanding that emerges from this inquiry is that, far from being an objective act of perception, the experience of listening is a mediated, hermeneutic, dialogic process and as such it requires special consideration.

Obviously, this insight is hardly original, nor is it new to psychoanalysis. In his renowned article on “evenly suspended attention”¹⁵⁴, Freud (1912) attempted to deal with precisely this point, suggesting a listening method that would insure the physician’s neutrality. However, in following decades, psychoanalytic voices coming from various theoretical approaches have gradually and consistently problematized the ideal of neutrality on which analytic listening was based. As I have argued, the interdisciplinary dialogue increasingly maintained between psychoanalysis, philosophy, linguistics, neuroscience and development has questioned the transparency of the process of listening, emphasizing its situatedness. The outcome of these developments is that the complexity of psychoanalytic listening as both a form of action and a professional stance is increasingly being recognized. Nevertheless, some of its aspects await sufficient elucidation, especially those experiential, embodied aspects of listening. These will be the main concern of the present chapter.

As a point of entry to these matters, I discuss analytic listening as it appears in the literature. Later on, through a discussion of the use of musical metaphors in psychoanalytic discourse, I propose a “musical” listening for the psychoanalytic

¹⁵⁴ I will elaborate this concept in the next section.
situation. The relevance of music to illuminate the psychoanalytic clinical situation is related centrally to its ability to represent emotions in its own terms, as elucidated by the aesthetician Suzanne Langer (1970) and the music philosopher Victor Zuckerkandl (1957). In conclusion, the second part of this chapter engages the auditory experience of sound and music as a unique phenomenon that is of particular value to psychoanalytic listening. Following the ideas of Langer and Zuckerkandl, I argue that a musical consideration of listening in psychoanalysis opens up new notions of time and space, subjectivity and intersubjectivity. Through these notions, the musical experience emerges, far beyond its function as an aesthetic object, as a body of knowledge about a fluid, participatory and interpenetrating way of being-in-the-world, thus providing psychoanalysis with a powerful metaphor for the articulation of experience.

IV.2 Psychoanalytic Listening, Past and Present

The process of analytic listening is a unique form of enquiry and communication that does not exist as such in any other human exchange. Analytic listening is taken to be the *sine qua non* of psychoanalysis, considered by many professionals as the most basic analytic skill, the essence of the analytic process, the very framework that the analyst provides for the patient (Freedman, 1983, Makari and Shapiro, 1993; Meissner, 2000). Stein (1999) claims that the fundamental importance of listening in psychoanalysis transcends and unifies all other components of theory and technique: “No matter an analyst’s theoretical orientation, and irrespective of the conceptual frame in which the listening is organized or of its focus at any given moment during
an analytic session, listening is an activity underlying every aspect of the analytic interaction occurring within the auditory sphere” (Stein, 1999, p.386). A stance, a skill and a framework all at the same time, listening pervades all aspects of psychoanalysis, from the setting of the analytic stage to its *dramatis personae* and their role: the analytic couch that reinforces speech and audition; the deliberate asymmetry in the analytic pair’s roles, whereby the patient mostly talks and the analyst mostly listens; the fundamental psychoanalytic rule that, through free association, generates psychoanalytic “material” as object of listening. Finally, listening is both the foundation and the epitomy of the analytic project of making the unconscious conscious.

Despite its omnipresence, psychoanalytic listening received relatively little attention in the literature, a fact readily acknowledged by various authors though not sufficiently questioned (see Bohm, 2002; Epstein, 1988; Gray, 1973; see especially Makari and Shapiro, 1993, for a succinct history of analytic listening). The transparency of analytic listening is evident in the very ideal of scientific neutrality pursued in early psychoanalysis, as its very nature as a methodical tool of data gathering and observation was not thoroughly explored. Freud addressed it explicitly in only one technical paper, in which he formulated the radical idea of “evenly suspended attention” as scientific stance and method (1912). Through this form of attention, Freud sought the suspension of criticism and preconception and the placing of differential focus and/or invested interests in the patient. But this technique, deemed by Freud as simple and “sufficient for all requirements during the

---
155 In fact, Freud dealt with the concept of evenly suspended attention in papers ranging from 1900 to 1923 (in Epstein, 1988), but addressed it directly only in this single paper.
treatment” (ibid. p. 111), was soon enough problematized. Freud himself recognized early on that the flow of free associations is affected by the interplay between wish and defence and admonished that the analyst should pay equal attention to each of the three aspects of intrapsychic conflict, namely the ego, superego, and id (Brenner, 2000). Consequently, the idea of evenly hovering attention as an exclusive listening technique was modified by several theorists. Ferenczi debated the shifting between the analyst’s “free play of association and phantasy, the full indulgence of his own unconscious” and a stance of logical scrutiny (quoted in Makari and Shapiro, 1993). In his book “Listening with the Third Ear”, Theodore Reik (1958) famously discussed analytic listening. He rejected the idea of evenly suspended attention and proposed instead “freely floating attention”, adding that psychoanalytic listening is a form of alert listening that always involves oscillation, continually shifting between a global, general reception of stimuli, and a more focused and directed attention (as I argue in the next section, oscillation is indeed fundamental to listening, shifting between not merely two, but multiple objects and modes of attention). A different concept of listening emerged through the thinking of Kohut (1959), who proposed that psychoanalytic listening enabled an imaginary attunement to the mental position of the patient. He claimed that empathy is the exclusive way by which an analyst might know the inner state of the patient and the concept of empathic listening and vicarious introspection became central to the analytic movement of self-psychology (see especially Schwaber, 1983, 1984, 1986). The concept of empathic listening is particularly germane to the present study, as it emphasizes the affective, participatory aspects of listening, which come about largely through the nonlinguistic vocal-auditory dimension.
Although the debate on the nature of analytic listening steadily acquired impetus, it appears that the main interest was essentially directed at the various objects of listening (the patient's narrative, dreams and free associations, the analyst's countertransference), rather than to the investigation of the listening process itself. However, as we have seen throughout this study, the vicissitudes and developments of psychoanalytic theory have, over the ensuing decades, problematized traditional psychoanalytic tenets. In terms of listening, the conception of the analyst qua listener gradually changed from a neutral external observer to an active, permeable participant, inevitably implicated in the psychological processes unfolding in the transferential scenario. The gradual unsettling of the initial philosophic notions of neutrality, truth, the stability of internal or external reality, the challenging of the binary concept of psyche-soma, the growing emphasis on the role of language in the construction of self as culturally and socially situated and the debate regarding the ontological status of the subject have all deeply affected analytic listening (see Frie and Reis, 2001, for an overview of philosophical influences on psychoanalytic conceptions).

Through these recent developments, the literature on listening has seen a relative growth, whose thematic diversity indicates that the subject is becoming a matter of increasing analytic concern (see, for example, recent texts by Bohm, 2002; Carlson, 2002; Gamelgaard, 1998; Helm, 2000; Lothane, 2006; Makari and Shapiro, 1993; Meissner, 2000; Pine 2001; Spence, 1984; Stein, 2007). Various aspects of listening have been debated in the literature, the most recurrent themes being: elaboration and critique of Freud’s idea of evenly suspended attention (e.g. Brenner, 2000; Epstein, 1988; Gardner, 1991; Pine, 2001); the concept of oscillating attention
(Ferenczi, Brenneis, 1994; Carlson, 2002; Freedman, 1983; Gardner, 1991; Helm, 2000; Reik, 1958); the influence of theory on practice (Bohm, 2002; Killingmo, 1999; Pine, 2001; Spence, 1984); the relation between listening and interpreting (Baranger, 1993; Faimberg, 1996); the nature of the listening process (Freedman, 1983; Makari and Shapiro, 1993; Meissner, 2000; Rizzuto, 1995, 2002; Smith, 1995); listening and its relationship to language and metaphor (Canestri, 2000; Gamelgaard, 1998); the listening stance (Epstein, 1988; Rubin, 1985; Ogden, 1997, 1999a, 1999b); empathy as a form of listening (Schwaber, 1983, 1984); intersubjectivity and listening (Beebe, 2004; Beebe and Lachmann, 2002; Frederickson, 1986; Frie and Reis, 2001). It is beyond the scope of this study to discuss these themes in full and, for our present purposes, it will suffice to note that this literature attests to the growing recognition of the complex and multidimensional nature of listening.

One issue of concern in the present study is that, all in all, most of the analytic debate regarding listening revolves around language. Expectedly, the “talking cure” correlates with a linguistic form of listening. Needless to say, however, analytic listening entails more than the decoding of linguistic meanings. Listening is at once a felt-experience, an interactive event and a mental act involving perceptual, cognitive, emotional and interpretive processes, all of which are directed at the totality of the patient’s utterance (which involves linguistic-discursive as well as nonlinguistic-affective expressions). One of the central criticisms voiced in the literature is that psychoanalysis has not sufficiently addressed (has not sufficiently listened to) the embodied, performative, nonverbal aspect of analytic interaction (Anderson, 2008; Aron and Anderson, 1998; Bucci, 1985, 2001; Canestri, 2000; Frederickson, 1986; Frie, 1999, 2002, 2003; Jacobs, 2001; Muller and Tillman, 2007;
Pally, 1996, 2001; Rose, 1980, 2004). This includes, as I argue, the nonlinguistic sonic
dimension.

The question is, then, how does one listen to this dimension, and, furthermore, how
does one conceptualize this form of listening? In an attempt to meet this challenge,
and through focusing on the nonverbal, embodied, sensual aspects of listening, I
propose in the following a mode of listening to the psychoanalytic clinical situation
that draws from the experience of musical listening.

IV.3 The Musical Nature of the Auditory Experience in the Clinical situation

IV.3.1 Music as Metaphor in Psychoanalytic Discourse

What does it mean to listen musically to the voices of the analytical exchange? A
short digression, elaborating the subject of music as metaphor, may provide some
answers. As I have argued elsewhere (Frank-Schwebel 2007), the metaphor of music
has been used increasingly in psychoanalytic literature to describe various aspects of
analytic experience. Over the centuries, music been widely employed in a
metaphoric capacity especially in the literal arts, i.e. poetry and literature, and
clearly became pertinent to them. However, its appearance in psychoanalytic
discourse is not as obvious. As I claim, it may indicate a new mode of analytic
thinking about the clinical situation. Nevertheless, in most cases the musical
metaphor does not extend beyond a vague poetic intuition and very few authors
have investigated the deeper relationship and commensurability of these two areas.
Through these metaphors, analytic experience was depicted as music (Bass, 2000; Ogden, 1996; Rose, 2004; Stein, 1999, 2007). Bass, for example, writes about a seminar he leads at the Institute for Contemporary Psychotherapy in New York called “Listening”: “I regard this seminar as a special sort of music appreciation class in which the music that we immerse ourselves in the sounds and silences of psychoanalysis itself” (Bass, 2000, p. 876). Ogden (1996) begins an article on technique with the following: “Debussy felt that music is the space between the notes. Something similar might be said of psychoanalysis. Between the notes of the spoken words constituting the analytic dialogue are the reveries of the analyst and analysand. It is in this space occupied by the interplay of reveries that one finds the music of psychoanalysis.” (Ogden, 1996, p. 882). In a similar vein, analytic interaction has been conceived as musical improvisation (Benjamin, 2002, Knoblauch, 2000). As described in chapter III, the musical concepts of resonance and attunement (Stern, 1985) have recently entered current psychoanalytic terminology, qualifying interpersonal emotional gesture and intent.

Several writers have specifically put forth an analogy between analytic listening and musical listening and helped elucidate the link between music and psychoanalysis, each to a different extent (see for example Bollas, 2007; Knoblauch, 2000; Lachmann, 2001; Nass, 1971; Ogden, 1999a; Rose, 2004; Stein, 1999; 2007). Knoblauch (2000), for example, devotes an entire book to the theme of listening to the “musical edge” of analytic interaction, through what he calls “process contours” – the moment-to-moment patterning and shifts in rhythm, tone and turn taking. Listening to the musical edge, according to Knoblauch, means attending to the
multiplicity of “pulses” of exchange, or the many languages, verbal and nonverbal, constituting communication. Through a close reading of a poem by Frost and a detailed discussion of an analytic session, Ogden (1999a) demonstrates how his analytic listening is aimed not at what lies behind the patient’s words, but at the sound and feel of the analytic happenings unfolding in the moment. Ogden places much emphasis on developing a “capacity to use language that does justice not only to the task of understanding and interpreting the conscious and unconscious meanings of patients’ experience; in addition, [the] use of language must be equal to the task of capturing and conveying in words a sense of what it is ‘that's going on here’ in the intrapsychic and intersubjective life of the analysis, the ‘music of what happens’ in the analytic relationship.” (Ogden, 1999a, p. 978-979, italics added).

Stein (1999) critically examines listening in music and in psychoanalysis, in an attempt to expand the understanding of the nature of the relationship between unconscious processes and the ways in which the “evanescent discourse between analyst and analysand is heard”. Stein views these two seemingly different listenings as having aesthetic, philosophical and psychological links, mutually resonating in two common areas, the one being engaged, attuned listening and the other - the communication of feeling through sound. He maps the aesthetic-emotional experience of musical listening onto the analytic scenario, stating that “because music is feeling, not sound, it is pure emotion, and presents a way of hearing and listening which can (and perhaps should) be a model of psychoanalytic intersubjectivity” (ibid. p.412). Stein further articulates his own clinical experience in musical terms (2007). Attending to the patient’s expressions, he wonders: “What is it to experience a relationship that sounds like this?” “What about the past, or this...
individual’s inner life, is being conveyed by how he or she sounds?” “Whom (or what) am I hearing?” (p. 60). In a recent book on music and psychoanalysis, drawing from aesthetic, neurobiological and psychoanalytic sources, Rose (2004, see also 1993) similarly discusses musical listening as resonating with analytic listening. In accordance with Langer’s musical theory (1970), he regards music as an audible representation of the emotional quality of subjective, lived time and investigates both the concrete and metaphoric links between this claim and the analytic situation. Like most of the authors discussed (see also Nass, 1973; Noy, etc.), Rose describes how he listens for the “‘music’ – i.e. the feeling – behind others’ speech and behaviour” (p. xxviii). For Rose, analytic listening interweaves cognitive and sentient-aesthetic modes of perception. To restate in sonic terms, he considers the temporal, nonverbal vocal and behavioural manifestations of discourse as they unfold in psychoanalytic time, giving audible manifestation, through dynamic sonic shapes, to the analytic partners’ “thoughtful feelings and feelingful thoughts” (ibid.).

Upon reading these highly dense, evocative texts, one cannot but be impressed by the effort these writers take in their attempt to integrate their sentient auditory experiences in music and psychoanalysis into new symbolic articulations. One theme that seems to be common to these writers is their sensitivity and attention to the (albeit evanescent) signifying value of nonlinguistic happenings occurring in the here and now, or to the “present moment” (Stern, 2004) of the analytic interaction as embodying affective subjective and intersubjective meanings. Their metaphoric musings seem to hit a right note, effectively evoking what transpires “between the notes of the spoken words” of the ineffable analytic experience.
How can the recent appearance of musical metaphor in psychoanalytic discourse be explained, and what does it suggest about a musical listening in psychoanalysis? The general literature on metaphor (see for example Black, 1962; Lakoff and Johnson, 1980; Langer, 1970; Noy, 1999; Sandbank, 2002) argues that new metaphors in language are created when the discourse lacks words to represent emergent perceptions and intuitions that are not yet conceptualized. Psychoanalytic discourse has similarly understood the pervasive use of metaphors in language and their source in somatic experience and affect (see for example Modell, 1997). Metaphors are the outcome of an imaginative mapping between two different domains, (in this case, music and analytic experience), revealing and at the same time, creatively conceptualizing new experiential horizons that were not yet labelled by language. This creative apperception, starting out as a transfer of meaning from one domain to the other, juxtaposing and thus prereflectively evoking their similarities and dissimilarities, is a form of imaginative play, through which new meanings are generated.

Through this argumentation, it can be assumed that the emergence of music as a common metaphor in analytic discourse is caused by an emergent need, shared by these authors, to describe and qualify certain aspects of analytic experience that, although sensed prereflectively, cannot be accounted for through ordinary linguistic means. Music provides these authors with a new metaphor to represent this experience. But although this analogy is explained in depth by some authors, in many cases it remains a mere poetic device. The next section provides an attempt to further elucidate this metaphoric phenomenon, which, as I argue, is significant.
Music not only helps to poetically conceptualize certain analytic intuitions, but radically illuminates the psychoanalytic clinical arena with a new light. Its eloquence as metaphor emerges from three sources:

1. Musical and verbal communications share developmental beginnings: mutual listening and vocalization are the fundamental avenues of communication out of which both verbal language and music branch out later on. As seen through developmental discourse, the musical metaphor has a material basis in early human existence. The music of communication and its link to innate human intersubjectivity has been considerably demonstrated by infant researchers. Malloch and Trevarthen (2009), for example, speak of “communicative musicality” as the basis of human companionship. They define musicality as the “expression of our human desire for cultural learning, our innate skill for moving, remembering and planning in sympathy with others that makes our appreciation and production of an endless variety of dramatic temporal narratives possible – whether those narratives consist of specific cultural forms of music, dance, poetry or ceremony; whether they are the universal narratives of a mother and her baby quietly conversing with one another; whether it is the wordless emotional and motivational narrative that sits beneath a conversation between two or more adults or between a teacher and a class... it is our common musicality that makes it possible for us to share time meaningfully together, in its emotional richness and its structural holding, and for us to participate with anticipation and recollection of pleasure in the ‘imitative arts’...” (Malloch and Trevarthen, 2009, p. 4-5). I readily include the psychoanalytic clinical
encounter as yet another instance of the dramatic, emotional temporal narrative inherent to the communicative musicality of human interaction.

In Western culture, language and music evolved and came to be conceived as separate fields, each with its own function and its own principles of operation. The analogy created by musical metaphors between music and speech works against the dichotomization (largely caused by culture’s visuo-logocentric tendencies) of these two sonorous activities. Through this metaphoric use, the music of speech and the signifying power of music are highlighted and relocated on a fluid continuum.

2. The musical metaphor reveals the structural affinities between the field of music and the psychoanalytic conversation, in which sound and time are the central elements that are interlaced in the transmission of the message. First, temporal experience is central to each realm. Time is embodied in speech and music in typical ways, materializing in different shapes and forms, involving memory, temporal sequence and continuity, repetition, expectation, and an immediate, lived sense of past, present and future. Second, both media generate meaning through the employment of sound and its components – each in its unique manner and for different purposes. While verbal language is discursive and propositional, music conveys affective and quasi-narrative meanings through its unique modes of signification (these issues have been addressed in chapter III). A third point which is particularly relevant to our subject and is further elaborated in the present chapter, is that both are live events that elicit an involved, participatory, sympathetic listening. In both, meaning is generated in the course of an active and interactive process of interpretation. And, as we have seen, both music and speech in
treatment are understood not only in relation to the present context. As situated events, they contain, within their very structure, implied listeners to whom their message is addressed. The musical metaphor thus calls forth the common roots and structure of both fields and repositions them on a fluid, multidimensional continuum, where their interpenetration and various forms of confluence become evident.

The common developmental source and the structural similarity of these two realms constitute the more obvious explanation for the eloquence of musical metaphors. Because of their similarity, the juxtaposition of the two domains enhances the musical nature of all utterance and attuned communication. It is perhaps meaningful that musical metaphors feature mostly in those strands of psychoanalysis (like the intersubjective and relational traditions) that privilege the curative power of the analytic relational experience itself. The musical metaphor in this discourse acts implicitly as a symbol of the curative effect of the attuned, empathic and intimate analytic relation as the foundation of the therapeutic experience. However, the appearance of the musical metaphor has further implications. At this point, I depart from the specificity of the metaphor and directly address music as an aesthetic and philosophic object that may expand our understanding of the analytic situation.

156 Such effects are centrally mobilized in the field of music therapy through actual musical activity – mainly by way of instrumental improvisation. Regrettably, the scope of the present study does not permit further elaboration of this issue, which greatly corroborates my present claims.
IV.3.2 Music as Symbolic of Emotional Experience

Music is deeply relevant to psychoanalysis not only by way of analogy or similarity, but as I argue, by illuminating it in more subversive ways that disrupt the well established analytic order and asserting its own hermeneutic value regarding analytic happenings.

One such way concerns the relation between language and experience. As we have seen above, the very structure of experience and the possibility of verbalizing it are at great odds. Music is an inherently experiential object that similarly resists verbalization. Thinking of psychoanalytic experience in a musical way problematizes the traditional exclusiveness of language in psychoanalysis, and makes the irreducible gap between language and experience tangible. Admittedly, this gap had always been acknowledged in psychoanalysis - at least, implicitly. Psychoanalytic treatment was not haphazardly constituted as a live, experiential encounter. Though its foundation is language, psychoanalysis cannot be conducted, for example, through written correspondence, and slips of tongue - those curious vocal mishaps - can only occur in real time. Nevertheless, experience was paradoxically disregarded by classical psychoanalysis as a “primary process” – an archaic mode of being that “needs to be outgrown” (Milner, 1987, p. 251) and transformed into the more mature, logical, symbolic, linguistic insight belonging to secondary process. It was not until analytic attention was directed at the analytic relation, at creative experience as something more than sublimation and at embodied experience as the basis of all symbolic activity, that the importance of “raw” experience began to be

By its very nature, when music is mapped onto the analytic situation, it exposes the gap between experience and verbal language, marking the centrality of lived experience, calling attention to its sensuousness, fluidity and intangibility. Yet, as we shall see, it also demonstrates its inherent intelligence and its different logic.

In this light, the use of musical metaphors seems to express the "unthought known" of the above authors (Bollas, 1987), signifying that analytic listening (as all listening), begins in an embodied, prerreflective, experiential mode. Linguistic, reflective and symbolic meanings generated in the course of listening are consequences of a primarily bodily, sensuous experience of sound, hence, listening must first be addressed through the effects that take place in the lived moment of the clinical encounter. Such listening entails, as Milner (1987) argued, a “concentration of the body”, or “body attention” (Gendlin, 2003) similarly refers to it as a bodily “focusing”). She writes:

What I mean by body attention or body concentration in the analyst is this: it is a state in which the direct proprioceptive body-self awareness ... becomes the foreground of one’s consciousness ... this kind of attention differs from the free-floating kind ... because it is not ‘in the air’: it deliberately attends to sinking itself down into a total internal body awareness, not seeking at all for correct interpretations, in fact not looking for ideas at all – although
interpretations may emerge from this state spontaneously. (Milner, 1987, pp. 239-240).

Milner further describes this form of attention as “wide focused. Its object, one’s own internal body perception, is inarticulate, dark, and undifferentiated.” (ibid.) This conception of listening echoes the phenomenological emphasis on embodied experience maintained throughout this study. Accordingly, our immediate awareness is prereflective and always inherently related to our bodily experience. As psychoanalyst Lewis Aron argued (1998), “the foundations of the capacity for self-reflexivity lie in our bodily sensations” (p. 23). Several other authors have also maintained (notably Frie, 2003; Gallagher and Zahavi, 2008; Zahavi, 2000, 2004, 2005) that it is precisely our bodily experience that forms the basis of subjective experience and enables reflection. Thus, linguistic articulation is a necessary, but not a sufficient means for the constitution and expression of subjectivity. Rather, as Aron proposes, we might think of it as being constructed by the ongoing interaction between language and the body.

In this vein, listening to the analytic discourse as a kind of music means surrendering to the “inarticulate, dark” sense and sensuousness of the sonic stream, where meanings are indeterminate and in an emergent state of becoming (thus French philosopher and musicologist Victor Jankelevitch [1983] phrased the eternal predicament of music). Listening musically to the occurrences of the analytic situation means abandoning oneself to the enveloping presence and texture of its sounds, giving in to their immediate emotional impact, to the ways they come towards us, touch us, invade us, carry us into unforeseen realms of meaning. It
means surrendering to the uncertainty of what is lived subjectively in the present moment but not yet cognitively known; yielding to the seemingly illogical logic of the auditory realm, to the “sound of sense”, to use Robert Frost’s sensitive phrasing, before arriving at (reflective) sense. Listening is thus part of a sentient interaction; a lived experience of continually affecting and being affected through sounds (see Rizzuto, 1995, 2002, on the relationship between sound and sense). Music, therefore, directs our awareness toward the nature of analytic experience and indicates that psychoanalytic inquiry is not carried out exclusively through the decoding of linguistic discourse, but also through a sensuous involvement and performance of speech and listening that unfolds temporally in an intimate, attuned, intersubjective encounter. In chapter III, through various musical concepts and principles, I demonstrated how, through this notion of musical listening, the sounds of patient and analyst as participants in a shared voyage can be attended to, and how their affective, interpenetrating discourse gradually generates new knowledge about the patient’s psychic life.

Yet, there is yet another important implication of the use of music in the psychoanalytic domain. Not only does it make us aware of the centrality of sentient experience, it also asserts its intelligent structure. Through musical experience and discourse, the stream of experience and perception is demonstrated to be neither archaic nor inchoate, but a polyphonic, multidimensional realm, continually striving towards organization, form and balance. Hence, at the same time as being fluid and indeterminate, music captures the quality of experience arguably better than any other art form. Langer (1970) has famously advocated the logical structure of
music, construing it as a symbolic medium due to its connotative relationship (through similarity of form, or isomorphism), with subjective experience.

Zuckerkandl (1957), though in different terms, similarly addressed musical experience as illuminative of the unique order of auditory space, elucidating its radical difference from visual space and asserting that auditory order and logic in music come about, phenomenally, through the immediate, prereflective experience of harmony, melody and polyphony as meaningful occurrences that unfold and coalesce in the moment-by-moment listening experience.

These contentions about music are significant in the present psychoanalytic context, especially because of the relation between music and affect. Langer proposed that music is fundamentally related to affect, providing insight into the morphology and structure of feeling and affective life and representing it in its unique manner.157 Music does so in a way that is intrinsic to the sonic medium and cannot be expressed through any other terms, being essentially untranslatable to other media, especially verbal language. Langer states that

logically, music has not the characteristic properties of language – separable terms with fixed connotation, and syntactical rules for deriving complex connotations without any loss to the constituent elements ... Yet it may be a

157 That music is related to affects is commonplace. However, the ways in which it conveys, arouses or represents feelings is a complex aesthetic issue debated throughout history. Most philosophers and aestheticians dealing with music, from Plato (1942) through St. Augustine (1962), Schopenhauer (1966), and, in contemporary discourse, Kivy (1980), Langer (1953, 1970) and Zuckerkandl (1957), to name just a few central representatives, have addressed the relation between music and affect and the related questions regarding its expression and representation through sonic means. As we have seen, music semiotics has debated the meaning of music and its affective power differently, seeing these as the outcome of the semiotic operation of the nonlinguistic sonic sign. The relation between music and affect has also been debated in psychoanalysis to a considerable extent (Feder et al., 1990, 1993; Ehrenzweig, 1965, 1967; Stein, 2004).
presentational symbol, and present emotive experience through global forms that are indivisible as the elements of chiaroscuro ... But it seems peculiarly hard for our literal minds to grasp the idea that anything can be known which cannot be named. Therefore philosophers and critics have repeatedly denied the musical symbolization of emotion. (1970, pp. 232-233).

Langer refutes this criticism of music, demonstrating the fallacy of logocentrism and positivistic thinking, which is based on the assumption that the rubrics established by language are absolute. According to her, it is the by virtue of the strength of musical expressiveness, that “music articulates forms which language cannot set forth [...] it is just because music has not the same terminology and pattern, that it lends itself to the revelation of non-scientific concepts” (ibid.).

Through the various phenomena that music embodies, Langer and Zuckerkandl thus claim the existence of a sonic logic and symbolic order that are fundamentally experiential and presentational – vastly different than the discursive or visual order – and are inherently linked to the realm of feeling. Beyond their importance in elucidating the artistic field of music, the ideas suggested by Langer and Zuckerkandl provide invaluable insights to human feelings and emotional experience.

Thus, these musical understandings can be mapped onto the psychoanalytic situation, in accordance with the assertion that verbal language is limited in its articulation of the affective experiences unfolding in the analytic encounter. Although Langer’s ideas are articulated in relation to the highly sophisticated medium of music, I argue that they can be applied to the analytic dialogue, essentially serving the same function – elucidating the expression and articulation of
emotion. Like musical experience, affective analytic experience is not discursively
told, but rather performed, undergone and lived in a sentient manner through
listening and the voice. Just as emotions are articulated and represented in music
through sound, so is the affective experience of the analytic partners articulated and
represented through their voices and reciprocal listening, thus constituting the
“music” of the psychoanalytic encounter. The transfer of musical knowledge to the
clinical situation facilitates a greater understanding of the emotive music of
conversation that unfolds beyond discursive meanings.

The idea of the “music of speech” is thus maintained to have more than metaphoric
value - by being solidly based on the concrete unfolding of the nonlinguistic vocal-
auditory dimension. Throughout this study, vocal manifestations of emotion were
considered not as direct symptoms or indices of a naturalistic reified interiority, but
as semiotic forms that are manipulated and stylized (though unconsciously and in a
lower level of complexity than in music), understood and interpreted - essentially
through prereflective, embodied consciousness. The referents of such semiotic vocal
forms are mostly emotions and energetic states as subjectively felt by the individual
and the interacting dyad. However, these vocal forms are not static representations
or labels of emotion (cognitively known as “categorical feelings”, see chapter III), but
are rather performative, presentational sonic forms of the quality and dynamics of
emotion as they unfold between conversational partners. Implicitly understood by
the listeners as significant, the nonlinguistic sonic dimension mobilizes a form of
knowing which is not inferior to discursive knowledge, nor is it archaic, primitive or
pathological. Rather, it is the embodied pool of experience from which all reflective and symbolic activity emerges. Langer states that

the assignment of meaning [to music] is a shifting, kaleidoscopic play, probably below the threshold of consciousness, certainly outside the pale of discursive thinking. The imagination that responds to music is personal and associative and logical, tinged with affect, tinged with bodily rhythm, tinged with dream, but concerned with a wealth of formulations for its wealth of wordless knowledge, its whole knowledge of emotional and organic experience, of vital impulse, balance, conflict, the ways of living and dying and feeling. (p. 244).

To paraphrase Langer, nonlinguistic sonic forms are recognized as significant by the listeners’ prereflective consciousness. Although these forms are not symbolic in the linguistic sense, they are nevertheless representational, or quasi-symbolic. Various authors have addressed them using different terminologies: Langer called the musical symbol an “unconsummated symbol”, pointing out its dynamism, ambivalence, and context-dependence; music semioticians have designated them, through Peirce’s theory, as sonic icons and indexes (sometimes also as symbols, however symbols which retain their vivid iconicity and indexicality); Stern conceptualized them as “vitality affects” and “pre-narrative envelopes”, and Trevarthen has addressed them as nonverbal vocal narratives. All of these writers intend to portray these vocal happenings as having a form and a signifying power which operate differently than those of literal symbolism. Though they are not yet symbols in the full sense, they constitute various intermediate states of
representation and expression. In all of these conceptions, the dichotomy of verbal/nonverbal is rejected and a more fluid mode of symbolic expression is provided: a multidimensional model in which vocal events are organized and understood through various mental-affective bodily modes and various forms of symbolization, facilitating the nuanced representation of emotional complexity and ambivalence. In chapter III, I have argued that the music of speech is not a mere metaphoric conjecture, but it has a concrete manifestation in the unfolding of sonic elements, which can be perceived in a real, sentient way.

Two more points deserve attention: first, the conception of musical listening places central focus on the “present moment” (Stern, 2004) of the psychoanalytic encounter. The nonlinguistic sonic dimension, with the concrete realness of its sounds and its immediate emotional impact, happens in the now. Vocal-auditory experience occurring in the present moment thus offers a point of entry into the realm of subjectivity and intersubjectivity. This experiential approach is different from the theoretical, inferential models that characterize traditional psychoanalysis. It maintains that insight emerging from the nonlinguistic sonic dimension comes about phenomenologically, through lived-experience, performance and participation.

The second issue concerning musical listening is that it highlights analytic listening as a participatory activity. In this sense, it can be said that the analytic listener is not, as Western logic would have it, a detached bystander observing objects that are “out there” in space and decoding their meanings in a controlled manner. The musical auditory experience reveals the analytic listener (and here I include both therapist
and patient), as implicated, penetrated, carried away, enveloped and surrounded by the sounds of the other, as well as her own. Zuckerkandl (1957) succinctly describes this musical experience:

> Far from taking us out of space – as common opinion holds – music discloses to us a mode of being of spatiality that except through music, is accessible only with difficulty and indirectly. It is the space which, instead of consolidating the boundaries between within and without, obliterates them; space which does not stand over against me but with which I can be one; which permits encounter to be experienced as communication, not as distance; which I must apprehend not as universal place but as universal force. (Zuckerkandl, 1957; p. 339)

The concept of the investigator as a participant, affecting and affected by her field of investigation has been widely elaborated, anywhere from anthropology to quantum physics and, of course, in psychoanalysis. However, it seems to me that the idea of musical participation explains more accurately the nature of the deep involvement of the analytic listener. Participatory listening is even more strikingly evident in musical improvisation, in which musicians not only play their instruments, but, constantly and simultaneously listen (again, in an embodied and prereflective manner). In the course of musical action and decision-making, they simultaneously perceive, attend to and process a multitude of aspects of the musical event (for example the other player’s sound or their own, the general “groove” or the mutual interaction). Ethnomusicologists Keil and Feld (2005) speak of musical participation, to which they also refer as “grooving”, as a sense of unified relatedness and
experience, as a “sharing of the other’s flux of experience in inner time, thus living through a vivid present in common constitutes the mutual tuning-in relationship, the experience of the ‘we’” (quoting Schutz, ibid., p. 129). Keil is vehement about the importance of participation, claiming that

the grooving of two tennis players or fencers is similar to the interplay of a jazz rhythm section at some level. That kind of physical grooving, being together and tuning up to somebody else’s sense of time is what we’re here on the planet for ... because of the way science and rationalism and empiricism have tended to squeeze the participatory out from everybody’s lives, it’s all the more important that we treasure the musical experience for keeping it there (ibid., p. 24).

Keil’s words imply that all social actors (like tennis players, musical improvisers, and psychoanalytic partners) are simultaneously also listeners who, through their mutual listening, continually seek correspondence, attunement, and meaning.

This conception of participation decenters the linguistic listening model characteristic of Western modes of conversation, with its (reified) distinct, sequential turn-taking, where speaking neatly alternates with listening. The structure of turn-taking, though obviously important in facilitating the clear and distinct flow of information, obfuscates the very participation which blurs the imagined boundaries existing between participants. Musical participation also hints at the simultaneous, fluid mode of acting and listening which takes place through verbal conversation, as the speakers’ sounds, rhythms and silences interpenetrate and create a shared

---

158 See Feld’s (2005) fascinating account of a different, polyphonic and multimodal structure of communication of the Kaluli people of Papua New Guinea.
texture. This conception of listening also makes it evident that the analyst is not the only listener in the room, and that the patient is also a sensitive listener to the analyst. Very little analytic literature addresses this specific point (exceptions are for example Aron, 1996; Faimberg, 1999; Meissner, 2000), as well as the participants’ listening to their own, inner and outer voice (but see Ogden, 1999a). These issues constitute important areas for further research.

Conceiving of listening as musical and participatory helps understand certain interpersonal phenomena and the manners in which they are manifested through the analytic partners’ vocal-auditory interpenetration. As discussed in chapter III, Stern (1985) and Trevarthen (1999-2000) have demonstrated how intersubjective states are vocally and temporally constituted and how affect attunement occurs through such musical parameters. Conversely, musical theoretical discourse addressed the nature of the emotional identification and sympathetic resonance of listeners with musical works (for example, HaCohen, 2001b, and as we have seen, Langer, 1970). In chapter three, I have suggested the phenomena of vocal duet and vocal ambience in psychoanalysis as constituted by participation and interpenetration.

Using the very same concept of musical auditory space I have portrayed above, psychoanalyst Jon Frederickson (1986) has attempted to explain the mechanism of projective identification. According to Frederickson, phenomenological theorists of musical experience have arrived at formulations which are quite similar to the psychoanalytic understanding of projective identification. Quoting various authors, he describes projective identification as a
mode of communication by which the projector makes himself understood by (unconsciously) exerting pressure on the recipient to experience a set of feelings similar to his own ... Just as we find in projective identification, music exerts a type of pressure on the listener ... Music, the thing ‘out there,’ although acting at a distance is in immediate contact with the listener’s mind and body. He undergoes its meaning. Music intrudes itself, denies the distance and enters into him. (Frederickson, 1986, p. 647).

As we have seen in the previous chapter, because of the peculiarities of auditory space, disavowed or repressed experiences and feelings can be projected in a nonverbal vocal manner onto the recipient, drawing him to (unconsciously) undergo their hidden meanings. These considerations demonstrate that intersubjectivity, empathy, projective identification and other dyadic interpersonal psychological phenomena are neither speculative nor mysterious processes, but unfold subtly through the vocal-auditory continual participation and interpenetration of the analytic pair.  

IV. 4 Listening - in a Manner of Conclusion

In the present study, I have attempted to make the sounds of the psychoanalytic encounter heard in a different way. My starting point was the idea that the sounds of speech in the clinical conversation coalesce, beyond discursive meanings, into a nonlinguistic sonic dimension that is vital to the psychoanalytic enterprise. I maintained that, despite the fact that psychoanalytic practice relies inherently on

Accordingly, although they are not discussed in the present context, this obviously occurs also through gestural and facial interaction.
voice and audition, it had, for various cultural and epistemological reasons, turned a deaf ear to this sonic dimension of signification. My main goal was thus to retrieve the nonlinguistic vocal-auditory dimension from the margins of analytic discourse and elucidate its signification and modes of operation. This required a strategy that would facilitate a new, different way of listening to the sounds of the psychoanalytic situation.

A central move was to counteract the binary pair verbal/nonverbal and reposition sound in a more fluid context, linking sound in human development, music and speech, demonstrating their common roots and operational principles as well as their divergences. For this purpose, I turned to disciplines that have significantly dealt with sonic experience in various capacities, centrally music aesthetics and semiotics and contemporary scholarship on development and subjectivity. My theoretical compass integrated a phenomenological stance and method with a Peircean conception of the musical sign and the Bakhtinian dialogic theory of human utterance and interaction. As argued in chapter II, these different theoretical frameworks were found commensurable in their shared focus on vital issues: the centrality of the body in phenomenal, lived experience; the view of communication as semiotic activity and process (rather than language as structure); the dialogic, intersubjective matrix that constitutes the subject; and the retrieval of the subject’s uniqueness and agentic power as expressed through the voice.

The investigation of the sonic dimension generated a multidimensional model of the voice and of listening. In this model, human utterance was understood as a nexus of various interacting forces: body and culture, language and music, self and other.
Linguistic and nonlinguistic aspects of the voice were conceived as continuous, rather than dichotomous, dimensions in communication. However, it was suggested that it is mainly the nonlinguistic voice and its embodied, performative and addressive impulse that infuses the speech-event with a sense of presence and vitality, and the impartiality of language - with the speakers' own voice.

Through this conception of the sonic dimension, I articulated several phenomena that included vocal principles of operation as well as specific vocal forms of varying complexity. These were claimed to operate in a manner similar to that of a musical system, being organized hierarchically and systematically through principles resembling those operating in music, representing the morphology of affective experience in similar ways. The totality of phenomena encompassed by this investigation was viewed as emerging from, and at the same time articulating, a dynamic, dyadic, intersubjective context. Much of the sonic dyadic interaction was understood through an analogy to dyadic musical improvisation.

The consideration of sound in the present study has led to a new conception of listening and the voice than those proposed by visuo-linguistic accounts. Nonlinguistic vocal expression was suggested to be a unique mode of affective expression that exceeds linguistic and paralinguistic meanings. Thus, the nonlinguistic voice was argued to be a central means for the expression of individuality and agency and for the articulation of relational meanings in psychoanalytic interaction. In the relational context, my conception of the nonlinguistic sonic dimension revealed, in a material way, the fluid interpenetration of participants in the intersubjective field. This renewed awareness of the vocal-
auditory space was suggested to have far reaching implications regarding certain established analytic conceptions and well known dichotomies. The concepts of oscillating attention and participation were particularly revelatory of the diffuseness of boundaries and horizons and the interpenetration and simultaneity of phenomena, as opposed to the relative solidity and distinctness characterizing the visual sphere. They demonstrated the multidimensionality of auditory space, where events, phenomena and the various layers thereof “bleed” into each other, simultaneously appearing and dissolving in space; where the clear-cut distinction between inner and outer, past and present, speaker and listener, subject and object, is dissipated. The deeper significance of musical listening is that it subverts the ideals of physical space and its attending notions of determinacy, separateness and exclusivity of meaning, which govern Western thought and language, marking them as illusory, highlighting the contextuality and transience of experience and raising fundamental questions about the stability of phenomena. However, and in contrast to poststructuralist views, the agency and self-authoring of the uttering individual was maintained and argued to be continually constituted through the unique operation of the nonlinguistic voice.

The nature of the present investigation demanded an intense focus on the nonlinguistic sonic realm. As I have maintained, this focus was only methodological. Human utterance is ultimately viewed as an entity that represents myriad blends of various levels of expression and representation. The interrelation between these levels requires further elucidation than the scope of the present study permits. A further interdisciplinary dialogue with music seems especially promising for the
investigation of this issue, especially since music has produced various mixtures of nonlinguistic and linguistic forms (such as songs, opera and Lieder). Such texted musical forms might serve as an important source for the further clarification of the interaction between these diversified modes of signification.

The present investigation, though seemingly circumscribed to listening and the voice in psychoanalysis, has touched upon fundamental issues such as the relationship between language and lived experience, the role of the voice as a signifier of the body and its relation to affect, subjectivity, intersubjectivity and agency. Human utterance was posited as simultaneously containing various signifying capacities (such as symbolic, iconic and indexical), raising essential questions about symbolism and representation, especially regarding emotions. Additional questions arose regarding the nature of the interplay between prereflective attention and consciousness, and their relation to psychoanalytic conceptions of the unconscious. The vastness of these issues could only be demarcated in the present context, providing directions for future research.

The last word goes to music and its contribution to the conceptualization of experience and specifically psychoanalytic experience. In the course of the study, music emerged, beyond being an aesthetic object of experience, as a medium and body of knowledge that can contribute to the expansion of psychoanalytic perspectives in the investigation of human experience and its relation to language. Through music, striking insights were obtained about the dynamic unfolding of meaning in the relations between subject and object and between the inner and the outer world. In the introduction to his insightful book on music, Sound and Symbol,
Victor Zuckerkandl poses: “What must the world be like, what must I be like, if between me and the world the phenomenon of music can occur? How must I consider the world, how must I consider myself, if I am to understand the reality of music?” (1957, p.7).

Zuckerkandl takes music to be a phenomenon in the world, reflecting a reality that is, however, not self evident. In that spirit, I have endeavored, through the reality of music and the sensuous tangibility of sonic experience, to better understand the psychoanalytic world. Rather than seeing music as external to the analytic project, I have positioned it in the analytic scenario as a voice that dialogues with psychoanalytic happenings. Far from being an intellectual exercise, this step stemmed from my own powerful personal and professional experiences of music as a medium of expression, communication and therapy. My musical intuitions and practical experiences, however, had to find conceptual formulations through words. As rewarding an endeavor as this proved to be, formulating and conceptualizing these experiences was perhaps the most frustrating and difficult aspect of the research – echoing in many ways the theme of the study itself: the attempt to create a dialogue between experience and concept, between sound and word. In many ways, this attempt at dialogue also parallels the central question and purpose of the present study: the dialogization of music and psychoanalytic practice, of embodied experience and reflective thought, based on a conviction that human utterance contains not just a proposition, but a mode of being and co-being with others and with the world that comes about musically. Much of my conceptualization is metaphoric, the body being its main reference. The articulations made in this study
avoided professional musical terminology and were purposefully designed to be understood as part and parcel of the ordinary, lived-experience that unfolds between mothers and babies, companions and conversational partners, and those engaged in other intersubjective voyages, like that of psychoanalysis. The methodical suspension of linguistic preconceptions and the enhancement of musical listening suggested here may thus provide an opportunity for perceiving the impact of our inherent musical knowledge of the world and our participation in it. Recognizing and acknowledging this impact may lead, in the psychoanalytic situation, to the reconciliation of dichotomous splits, thus to greater resonance between the various levels of human experience.
References


268


Bromberg, P.M. (1994). “Speak! that I may see you”: Some reflections on dissociation, reality and psychoanalytic listening. Psychoanalytic Dialogues, 4,517-547.


HaCohen, R. (2001a). Reflections and inflections of the transfiguring self: on the music of the ‘unheimliche’ in the works of Schonberg and Mahler. In *Österreich-


Streeter, E., (1999b). Definition and use of the musical transference relationship. In


Tolbert, E. (2001). The enigma of music, the voice of reason: "music," "language,"


