ISTANBUL BILGI UNIVERSITY
INSTITUTE OF SOCIAL SCIENCES
CLINICAL PSYCHOLOGY MASTER'S DEGREE PROGRAM

MUSIC OF THE TALKING CURE: EXAMINING THE INTERACTION BETWEEN PATIENT AND THERAPIST IN THE SPHERE OF SOUND

SİNAN TINAR
112629010

SİBEL HALFON, FACULTY MEMBER, PhD

ISTANBUL
2019

Music of the Talking Cure: Examining the Interaction Between Patient and Therapist in the Sphere of Sound

Konuşma Kürünün Müziği: Danışan ile Terapist Arasındaki Etkileşimin İşitsel Boyutunun İncelenmesi

Sinan Tınar
112629010

Thesis Advisor:	Dr. Öğr. Üyesi Sibel Halfon İstanbul Bilgi Üniversitesi	.....
Jury Member:	Dr. Öğr. Üyesi Alev Çavdar Sideris İstanbul Bilgi Üniversitesi	.....
Jury Member:	Dr. Öğr. Üyesi Yasemin Sohtorik İlkmen Yeditepe Üniversitesi	.....

Date of Thesis Approval: 30.05.2019

Total Number of Pages: 83

Keywords (Turkish):
1) Sözsüz iletişim
2) Psikanalitik psikoterapi
3) Ses
4) Öz-düzenleme
5) Ortaklaşa düzenleme

Keywords (English):
1) Nonverbal communication
2) Psychoanalytic psychotherapy
3) Sound
4) Self-regulation
5) Mutual regulation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This study focuses on the auditory dimension of the nonverbal interaction between the patient and the therapist in the process of a psychoanalytic psychotherapy. It employs computerized measurement of three parameters that are related to vocal communication; rhythm, intensity and pitch. This study has four aims: 1) providing a sonic description of sessions; 2) exploring patient and therapist's self-regulation and 3) mutual-regulation with regard to the three parameters; and 4) providing qualitative interpretations of the findings with the help of verbal content of sessions. To do this; six sessions are selected from a one-year psychoanalytic psychotherapy done with a 17-year adolescent; representing early, middle and late phases of therapy. Through time-series analyses, the autocorrelation of patient and therapist and the cross-correlation between the patient and therapist on three parameters are analyzed. For each session, a transcript of verbal communications during a selected "critical moment" is presented to provide lexical content for qualitative analyses. Statistical analyses revealed significant autocorrelation of intensity, pitch and rhythm for the patient in most of the sessions. On the other hand, for the therapist, no such significant autocorrelations are found except for one session. Also, there was significant cross correlation for rhythm in five sessions, pitch in one session and intensity in one session. These findings suggested that throughout the therapy process, while the patient engaged in varying degrees of self-regulation of her own vocal states; the therapist mostly followed and matched the vocal rhythm of the patient. Beginning with the second half of the selected six sessions, the patient increasingly matched her rhythm, and in one session matched her pitch with the therapist's. The outcome of statistical analyses and descriptive results are used to interpret the verbal content, and vice versa. Observations are discussed from the vantage point of several psychoanalytic concepts. Overall, the early phase of the psychotherapy process is used to depict how therapist's mirroring can facilitate the regulation of patient's affect as well as opening a space to think; and how the introduction of new topics can have a procedural impact on the patient's vocal

rhythm. The middle phase illustrated how disruptive enactments and their subsequent repair can catalyze the emergence of new states; and the last phase showed how leaning toward mutual regulation can be an indication of a symbiotic merger fantasy in the face of upcoming separation.

Keywords: Non-verbal behavior, psychoanalytic psychotherapy, rhythm, intensity, pitch, mirroring, enactment, self-regulation, mutual regulation.

ÖZET

Bu çalışma, psikanalitik psikoterapi sürecinde danışan ile terapist arasındaki sözsüz iletişimin işitsel boyutuna odaklanmaktadır. Bu araştırmada, sesli iletişime ait üç değişkenin, ritim, sesin şiddeti ve sesin perdesinin (pitch), bilgisayar tarafından ölçülmesi ile 1) seansların işitsel açıdan betimlenmesi; 2) danışan ve terapistin davranışlarının öz-düzenleme (self-regulation) ve 3) ortaklaşa düzenleme (mutual regulation) düzenekleri açısından araştırılması; 4) elde edilen bulguların seanslardaki sözel içerikle birlikte ele alınarak niteliksel olarak yorumlanması gibi amaçlar hedeflenmektedir. Bunun için, 17 yaşındaki bir ergen ile sürdürülmüş, bir yıllık bir psikanalitik psikoterapi sürecinden, erken, orta ve geç dönemleri temsil edecek şekilde 6 adet seans seçildi. Zaman serileri oluşturularak bahsi geçen üç değişkene ait otokorelasyon ve çapraz korelasyon analizleri yapıldı. Her seans için, seçilmiş bir "kritik an" esnasındaki sözel etkileşimin deşifresi, kalitatif analizlerde kullanılması amacıyla sunuldu. Danışan için; ritim, ses şiddeti ve perdesi açısından neredeyse tüm seanslarda istatiksel olarak anlamlı otokorelasyonlar bulunmasına rağmen, terapist için sadece bir seansta otokorelasyonlar bulundu. Ayrıca, tüm seanslarda ritim, bir seansta ses şiddeti ve bir seansta da ses perdesi açısından istatiksel açıdan anlamlı çapraz korelasyonlar olduğu gözlemlendi. Sonuçlar, terapi süreci boyunca danışanın ritim, ses şiddeti ve ses perdesi açısından düzenli olarak kendi kendisini düzenlemekte olduğunu işaret etmektedir. Öte yandan, terapist, neredeyse her seansta danışanın ses ritmini takip etmekte ve kendi ritmini danışanınkiyle eşleştirmektedir. Ayrıca, danışanın sürecin ortalarından itibaren, ve giderek artan bir şekilde, terapistin ritmini; bir seansta ses şiddetini ve bir seansta da ses perdesini izlediği ve kendi davranışlarını terapistinki ile eşleştirmekte olduğu gözlemlendi. Betimleyici ve istatiksel sonuçların seans içerikleri ile birlikte ele alınması ile süreç, çeşitli psikanalitik-klinik kavramlar açısından yorumlandı. Sürecin ilk evresi, terapistin aynalamalarının nasıl danışanın duygulanımını düzenlerken aynı zamanda düşünmek için bir alan açtığını; ve ortaya çıkan yeni konuların danışanın konuşma ritmi üzerinde nasıl bir etkisi olabileceğini göstermek için kullanıldı. Orta evrede, ilişkisel sahneye koymaların yıkıcı etkilerinin

onarılması ile nasıl yeni durumların ortaya çıkabileceği; son evrede ise danışanın ortaklaşa düzenleme süreçlerine yaslanmasının, yaklaşan ayrılık karşısında nasıl bir bütünleşme fantezisinin işareti olabileceği tartışıldı.

Anahtar kelimeler: Sözsüz iletişim, psikanalitik psikoterapi, ritim, ses şiddeti, ses perdesi, aynalama, sahneye koyma, öz-düzenleme, ortaklaşa düzenleme.

CHAPTER 1
INTRODUCTION

Most forms of psychotherapy descend from psychoanalysis, which was famously called by one of Sigmund Freud's patients as "the talking cure". Although the practice of psychotherapy basically consists of two or more people in a room speaking with each other, relatively little has been written regarding the sonic dimension of those verbal interactions. Sounds constitute the "scaffolding" that carries the meaning of the words (Beebe, Jaffe, Lachmann, & Feldstein, 2000, p. 101) and those sounds possess individual characteristics such as their loudness, pitch, or timbre. In addition, the sequential arrangement of sounds in time creates a flow having certain qualities. For example, linear arrangement of the intensity of sounds in time can create a crescendo, starting from quiet and getting louder over time; or a diminuendo, where the opposite process happens. Alternatively, a sequence of sounds having different pitch qualities creates a melody. Those melodies either conform to the rules of tonal harmony, or not bounded by the rules, drift into the realm of atonality. An arrangement of sounds or sound qualities can also create rhythms, that are either recurring, like the steady ticking of a clock; or they can be non-periodic, like the irregular rhythm of exploding fireworks. The larynx, or the voice box, is probably the first wind-instrument that humans ever had, and this instrument is being played every time someone is speaking, whether consciously or unconsciously. All in all, there is no "talking cure" without "organized sound".[1]

This study, in a global sense, is about the acoustic properties of speech that emerge during the process of a psychoanalytic psychotherapy. We hope to provide some knowledge regarding the following questions: What can the analysis of the dimension of sound offer for the scientific understanding of psychoanalytic therapies? How does patient and therapist respond to changes happening in each other's vocalizations? Do patient and therapist adapt to each other on the dimension

[1] Edgard Varèse, a 20th century avant-garde composer referred music as "organized sound" in order to extend the boundaries of classical definitions of music. Instead of a musician, he preferred to be called as "a worker in rhythms, frequencies, and intensities" (Varese & Wen-chung, 1966, p. 18).

of sound? Do they mirror each other? If they adapt their voices to each other, on which parameter does this adaptation is most visible? Is it rhythm, is it loudness, or is it pitch? How does the dyad's interaction progress in time? What can we learn about the dyad's relationship by investigating moments of convergence, divergence, or a moment of crescendo?

The musical qualities of conversations attracted theoreticians and researchers from a wide variety of disciplines; linguists, anthropologists, ethnomusicologists, infant researchers, psychoanalysts, to name a few. In the tradition of psychoanalysis and psychoanalytic psychotherapies, the interest towards the topic has increased steadily since the last century. At times, the topic is approached with a pure theoretical mindset. Arguments are usually based on clinicians' observations in their consulting rooms and elaborate descriptions of patients' qualities of speech are made in an attempt to stress the significance of those qualities. However, it is hard not to notice the scarcity of studies that combine these psychoanalytic perspectives with the evidence coming from empirical research. Moreover, in the tradition of empirical research, most studies seem to focus solely on one aspect of the conversations, such as rhythm. Also, a great number of empirical studies are conducted not in the context of the clinic but in experimental conditions. This study is an attempt to bridge the gap between theoretical and empirical by analyzing the vicissitudes of rhythm, intensity and pitch of patient and therapist's speech, throughout selected sessions from a psychotherapy done with an adolescent.

In the first section, we will review the literature, first by dividing the contributions into two threads, theoretical and empirical. Starting from Freud and post-Freudian thinkers we will review some of the major lines of thought. Then, we will present previous research done on rhythm, intensity and pitch. After that, we will do a summary and we will state how our study fills a gap in the literature. Following the statement of our exploratory hypotheses, we will present our method and results. Finally, we will discuss how our results relate to the previous empirical research and to clinical and theoretical points of view.

1.1 THEORETICAL

1.1.1. Freud and His Problem with Sound

In psychoanalytic practice, as Sigmund Freud envisioned it, although the treatment mainly depended on patient and analyst's speech, the focus was placed on the verbal contents of the speech rather than the acoustic properties. Freud regarded "verbalization" as the principal therapeutic activity, through which the patient's unconscious conflicts became visible to the analyst. By interpreting the hidden messages in the patient's discourse, the analyst made them conscious. The main mode of operation was lexical (based on meaning); patients' sentences were deconstructed, interpreted, re-contextualized and presented in a new light. What was missing in Freud's initial conceptualization of psychoanalytic practice is the non-verbal dimension (Jacobs, 1994). It is often stated that Freud had a problem with sound and music in general. Several authors have written about this topic in depth (see Barale & Minazzi, 2008; Cheshire, 1996; Nagel, 2013). One might argue that his general tendency to avoid phenomena that are "musical" have caused him to ignore, or at least not talk or write about the musical dimension of the psychoanalytic dialogue.

The idea that Freud having problem with music is mainly based on Freud's own statements. Although Freud showed a keen interest in some forms of art, such as sculpture or literature, he considered himself unskilled when it came to music. His words in the opening lines of *The Moses of Michelangelo* gives a hint about his thoughts on his experience of listening to music. He said;

> …with music, I am almost incapable of obtaining any pleasure. Some rationalistic, or perhaps analytic, turn of mind in me rebels against being moved by a thing without knowing why I am thus affected, and what it is that affects me. (Freud, 1914, p. 211)

In this passage, we see that Freud is placing emphasis on the emotional, evocative quality of the listening experience. He states that he is moved by the music in a way that he is rendered unable to grasp it with his reason. Similarly, In *The Interpretation of Dreams,* he mentions an occasion where he was humming a

tune from Mozart's *Marriage of Figaro* and said that "…it is a little doubtful whether anyone else would have recognized the tune…" (Freud, 1900, p. 207) This led some historians argue that he was "tone deaf" (Gay, 1988). Lastly, Barale and Minazzi (2008) have written about Freud's interest in the work of Theodor Lipps, whose ideas about unconscious processes had a great influence on Freud's thinking. They refer to a point where Freud "stopped" while reading Lipps. In Freud's words, he stopped:

> at 'sound relationships', (which) always vexed me because here I lack the most elementary knowledge, thanks to the atrophy of my acoustic sensibilities (Letter to Fliess, in Masson, 1985, p. 325).

On the other hand, some authors treating Freud's own statements as somewhat misleading, claimed that he was very much interested in musical matters. Diaz de Chumaceiro (1990) argues against the allegations that Freud was tone deaf. He states that it wouldn't be fair to diagnose Freud as tone deaf, since this condition is characterized by a total inability to perceive and recognize any melodic line. Freud's case was not that, although he might have had a difficulty in reproducing his favorite arias with his mouth. Diaz de Chumaceiro state that Freud enjoyed going to operas and had some close friends and colleagues who were deeply embedded in musical circles of Vienna.

It is argued that Freud's reluctance to talk about matters that are musical might be caused either due to some defensive maneuver or a deliberate decision on his part (Cheshire, 1996). Some explanations focused on biographical details that might explain such a defensive stance, such as his jealousy of his wife Martha's previous candidate in marriage who was a musician, his mother's passion in music, or the influence of him being a Jew in anti-Semitic musical Vienna. Barale and Minazzi (2008) don't find these biographical arguments convincing and propose several historico-cultural reasons that caused Freud deliberately distance psychoanalysis from sound and music.

Barale and Minazzi (2008) argue that music used to be central to science, but at the time of Freud, a split was happening between music as object of scientific inquiry, and music as the metaphor of the aesthetic experience. Music represented something that cannot be put to words – things that are only experienced intuitively.

In addition, music had strong connections with medicine, making use of its spiritualistic and suggestive aspects. It was also associated with hypnotism and mesmeric rituals. The idea of "being in tune with the universe" was prevalent and it is believed that pathologies are caused by "disharmonies". Barale and Minazzi state that "Freud's exclusion of the element of sound and music from the psychoanalytic edifice was inevitable… (and) …Freud was determined to cause that knowledge to grow in a soil sharply distinguished from a spiritualistic tradition" (p. 951).

Freud's exclusion of sound, as well as exclusion of the visual domain by the use of couch might have resulted in him not recognizing the potential values of the non-verbal domain that later psychoanalytic theoreticians insistently stressed. How was the Rat Man's tone of voice? In what way did he spoke to Freud when he was ruminating? How did Freud spoke to his patients when he gave interpretations, or asking questions? Did he produce sounds while listening? Did he match his voice to his patients' tone of voice? We don't know a lot since Freud didn't wrote about these topics. He probably gave his ear to his patients' voices in order to get information regarding their affective states, hesitations or censures in speech. Jacobs (1994) argues that he was indeed a good observer of non-verbal phenomena. He mentions Freud's depiction of one of his patients, Frau Emmy Von N., where he gives a description of her speech. In Freud's words, she spoke "in a low voice as though with difficulty and her speech was from time to time subject to spastic interruptions amounting to a stammer" (Breuer & Freud, 1895, p. 48-49).

1.1.2. Post-Freudian Interests

While Freud was reluctant talking about topics related to music and sound, let alone commenting on the musical dimension of the analytic dialogue, his contemporaries started to bridge the gap between psychoanalysis and music in various ways. Some authors wrote in an attempt to provide a psychoanalytic explanation of the enjoyment of listening music, while some wrote with the aim of

analyzing musical works and composers. More importantly, paralinguistic[2] features of speech also became a subject of interest, either attended as a means for diagnostic purposes or as a tool for recognizing the subtle changes of a person's internal dynamics. We will review some of the post-Freudian contributions while maintaining our focus on the ones that are related to the acoustic properties of speech.

Ferenczi (1932, as cited in Knoblauch, 2000) was one of the first psychoanalytic theoreticians that acknowledged the importance of non-verbal communication. Similarly, Karl Abraham included auditory factors in his *Sexualtheorie* (1914, as cited in Nagel, 2013). Isakower (1939) wrote about "the exceptional position of the auditory sphere" and developed the idea that the super-ego was a derivative of early auditory sensations. He even indicates that on Freud's initial graphic representations of the structure of the mind, a part of ego was called "Hörkappe", meaning cap of hearing. That part became the super-ego on Freud's later depictions. Max Graf, a close friend of Freud and a member of his famous Wednesday group, was a music writer and was the first one to apply Freud's ideas to the analysis of works by composers such as Mozart, Beethoven and Wagner (Graf, 1942 as cited in Nagel, 2013). Heinz Kohut (1950, 1957) wrote several papers on the nature of listening musical pieces from a psychoanalytic perspective. He emphasized the role of mastery. For Kohut, the experience of listening to music resulted in a regression to archaic, preverbal modes of psychic functioning and the enjoyment derived from such experiences is related to gaining mastery over that preverbal states. Greenson (1954) working with a patient who experienced the sensation of the sound "Mm" in his mouth during hypomanic episodes, became interested in this particular sound and traced its roots to the earliest feeding experiences. He also draws attention to the fact that words that have been used to signify "mother" in a lot of languages, started, or included this "Mm" sound.

It was probably Theodor Reik (1949, as cited in Templeman, 1977) who was the first to call attention to the musicality of a patient's speech. He suggested

[2] Paralanguage, is a term that is proposed by Trager (1958, as cited in Duncan, Rice, & Butler, 1968) and is used to refer to voice qualities associated with pitch, tempo, or intensity of speech as well as, pauses, sighs, disfluencies and other noises that don't have a lexical character.

that therapists should attend to patient's "particular pitch and timbre of voice, (their) particular speech rhythm… …variations of tone, pauses, and shifted accentuation" (p.11). He also likened the analyst's unconscious to a musical instrument which resonated with the patient's non-verbal messages. His idea of "listening with the third ear" encouraged analysts to be aware of the impact of the patients communications on their internal worlds (Safran, 2011).

Moses (1954, as cited in Bernholtz, 2013) wrote *The Voice of Neurosis* where he proposed that it would be possible to detect personality types by listening to the voices of patients. He believed that it was possible to discriminate schizophrenia from neurosis or detect whether someone is depressed or not, from a person's way of speaking. He used a technique that he called "creative hearing", focusing on the non-verbal qualities of sound, such as respiration, range, register, resonance, and rhythm (the 5 R's), alongside the verbal content (Bady, 1985; Templeman, 1977).

Silence has become an exceptional topic since the foundation of psychoanalysis (Arlow, 1961; Fuller & Crowther, 1998; Nacht, 1964; Reik, 1968; Sabbadini, 1991; Woo, 1999). It is usually regarded as a classical form of resistance, since the "fundamental rule" of psychoanalysis required patients to freely associate. Moments of silence were regarded as indicators of an internal conflict that is blocking the stream of talking. Fliess (1940), for example, distinguished between urethral, anal, and oral silences in an attempt to discriminate between types of silences regarding their purpose of use.

1.1.3. Shapes, Envelopes and Transformations

As followers of Freud began to extend the range of psychoanalysis to include the treatment of children, attention given to pre-verbal phenomena has increased, considering them an important factor influencing developmental processes. This increased attention seems to have been also fueled by the newly emerging therapeutic attitude being developed to treat severe types of psychopathology. Analysts needed to include non-verbal modes of intervening where their verbal interpretations started to appear limited. For instance, Anna

Freud (1959, as cited in Strean, 1969) argued that some patients had fixation points in the earliest, non-verbal stages of development and their treatment necessitated the application of non-verbal methods. Maiello (1995) stated that "the musical vertex of analytic attention is particularly important with psychotic and borderline patients, and in general in all cases where the attacks on linking make communication more difficult" (p. 40).

While the early skin experiences are understood as being an important factor on the formation of psychopathology (Bick, 1968) a cross fertilization between the topics of sound and skin seemed to occur. Niederland (1958) put forward the idea that the auditory experience is actually a contact experience. In a similar vein, Anzieu (1979, 1995) believed that contact with sounds was not so much different from the tactile experience. He proposed the idea of an audio-phonic skin, alongside his formulation of skins that are related to thermal, olfactory, and taste sensations. These and other components formed what he called the skin-ego.

Skin-ego is an interface between the internal world and outside world, separating the two, but also connecting them. It is an envelope that is wrapping up the baby. It is also a surface where the earliest interactions with the mother is experienced, such as the mirroring experience (Anzieu, 1995). Authors such as Winnicott (1967) and Lacan (1966) have been already stressing the idea of mirroring in their work. For Lacan, the "mirror stage" was a period where the child became fascinated with its reflection in the mirror and identified with it. This image provided the child with an organized, unified representation of himself. In Winnicott's ideas, it was the mother's face that provided the mirroring. Both theoreticians stressed the primacy of the visual domain. However, for Anzieu, earliest mirroring experiences happened in the auditory domain prior to the visual one. The way the infant is mirrored through sound, the way its mother is speaking with him was essential in the formation of the ego. Problems encountered in this process of acoustic mirroring were responsible for causing various types of psychopathology. For instance, Anzieu describes the voice of a mother of a schizophrenic as being "…monotonous (lacking rhythm) or metallic (lacking melody) or harsh (with a predominance of low-pitched sounds, which makes the hearer mix up the sounds and feel invaded by them)" (Anzieu, 1995, p. 226). He

also defined three characteristics of a "pathogenic sound mirror": its dissonance (when mother's vocal response doesn't match the nature of baby's needs and feelings), its abruptness (when the mother's voice changes instantly and rapidly), and its impersonality (when the mother's voice is devoid of a recognizable character). Alongside the function of mirroring, mother's voice also provided a holding function. The mother's incoming stream of vocalizations engulfed the baby, creating "a bath of sounds".

Maiello (1995) contributed to this line of thought by extending the relationship with sound towards the intra-uterine life. She stated that the proto-mental nucleus of the foetus "could use the mother's voice for the creation of an internal object with sound qualities, which in turn could become the ground on which rests the preconception of the breast" (Maiello, 1995, p. 27). For Maiello, the "sound-object" was the first object where the infant became acquainted with the experience of separation and loss. Therefore, the initial representations of object relatedness are first formed on the acoustic dimension of the intra uterine life.

Analysts working with the most primitive parts of psychic functioning developed an understanding of shapes and contours of sensory experience (Ogden, 1989; Tustin, 1984). Ogden, extending Klein's (1946) concept of the paranoid-schizoid position, proposed the idea of the autistic-contiguous position, as a more primitive non-symbolic mode of generating experience. In this mode, the form of the sensory experience, its rhythmicity, boundedness or its texture provided the subject with a foundation that later stages of psychic developments are built upon. In a similar vein, Stern (1998b, 2004) stressed the importance of temporal contours of behavior. For Stern, "vitality affects" had a different quality than other categorical affects, such as sadness or anger, and they can best be illustrated with dynamic-kinetic terms like "fading out", "bursting", "rushing", etc. Stern argues that starting from infancy, these "forms of feeling" play an indispensable role in our perception of the world. Furthermore, Stern identified three features of non-verbal behavior that could be matched between the mother and infant: its intensity, timing and shape. For Stern, matching these parameters formed the basis of affect attunement.

In the clinic, recent conceptualizations of psychoanalytic process depict the analyst as the receiver of the patient's unnameable psychic content. The analyst's function includes giving forms, shapes, and names to that unprocessed experience and facilitating the process of symbolization (Bion, 1970). The psychoanalytic space is seen as a reverberant place where the patient can remember the deepest forgotten melodies that are related to its past. It is also an emergent space that facilitates the discovery of new forms, melodies, rhythms, and ways of being with someone (Birksted-Breen, 2009; Erdem, 2005; Frederickson, 1986.; Grassi, 2014; Maiello, 2003; Ogden, 1989, 1999). In this conceptualization of the psychoanalytic process, the analyst might be likened to an interpreter of the music of the patient's unconscious. The analyst, attending to the sensory level, is capable of observing changes in rhythms, tones and affective contours. However, he or she is still understood as someone who transforms the incoming material for the patient, like a mother who distinguishes the baby's needs from the sound of her ambiguous cries. The analyst is not thought of as an active participant of the music that is being created in the room. In the next part, we'll present a line of thinking where the analyst-therapist eventually became the active participant of the process.

1.1.4. Relational Perspectives

Since the early 1980's, American psychoanalysis started to undergo "a relational turn". It was a critique of what they called a "one-person psychology", whose focus was solely on the intrapsychic conflicts of the patient. However, the "two -person psychology" they proposed, brought the relationship into the fore (Aron, 1990.; Ghent, 1989). This movement is obviously influenced by Winnicott and Sullivan's ideas who argued that an individual cannot be thought as isolated from the environment he or she is embedded (Mitchell & Black, 1996; Safran, 2012). So, the psychoanalytic practice eventually turned into a process where two people continuously interacted with each other on various symbolic and non-symbolic levels, whether they are conscious of these interactions or not.

This "relational turn" was fueled by evidence coming from empirical research on infants, attachment, and affective neuroscience (Safran, 2012). The

study of the infants necessitated the consideration of the pre-verbal modes of interaction. Methods that allowed the measurement of pre-verbal parameters have been developed. Similarly, with the emergence of new technologies, analysis of micro-moments became possible, tracking the moment-to-moment interaction of infant and mother, or patient and therapist.

The research of Daniel Stern (1998b) proposed a new relational infant, as opposed to the classical conceptualization of infants in psychoanalysis. Stern's infant was born with cognitive abilities and right from the start it was an active participant in the construction of the relationship. As the researchers ventured away from the categorical-verbal territories, new forms of "remembering" and "knowing" are conceptualized. Implicit memory and procedural knowledge were two newly emerging concepts that referred to phenomena that we know but which was beyond our verbal understanding. Procedural memories are "content-free, in the sense that they entail the learning of processes rather than information" (Beebe, Jaffe, Lachmann, & Feldstein, 2000, p.104). This procedural knowledge of being with oneself, and with others, formed in the earliest years of life, through the infant's interaction with its caregivers. The nature of those interactions was inevitably non-verbal, consisting of bodily movements, facial expressions, and voice. The non-verbal vocal interaction that is shared between the mother and the infant is one of the first experiences where the initial patterns of being with oneself and with others are constituted.

Within this vocal interaction, it is assumed that each member self-regulates its own states and also mutually-regulates each other. This was the combination of "one-person" and "two-person" views. Beebe et al. (2005), saw the mother-infant dyad as a system of "shared organizational forms" (p.25). They both had self-regulatory functions (like adapting to a change in stimulus) and they influenced each other continuously in a bi-directional way (but not necessarily symmetrical – the mother had greater range, control and flexibility). Beebe and Lachmann (2002) stressed the importance of attaining a "good enough" balance between self and mutual regulation processes in the treatment of adult patients. In their conceptualization, a fixation on either form of regulation indicated a type of pathological formation. Therefore, the therapist's objective is to create a balance

between the two forms of regulatory activity. Lachmann and Beebe (1996) also identified three organizing principles that guided their technique in adult treatment: ongoing regulations, disruption and repair, and heightened affective moments. Ongoing regulations happen throughout the psychoanalytic process and they shape the patient's expectations. Within the classical psychoanalytic tradition, the interpretation of unconscious material is regarded as the principal mutative factor. However, here, the sheer act of analyst's listening or questioning can also have a procedural impact on the patient's internal representations. Disruptions happening in the relationship between the patient and therapist are seen as inevitable and carefully attending to those moments can facilitate therapeutic progress. Lastly, Lachmann and Beebe argued that heightened affective moments open up the possibility of shaping the patient's self-regulatory style and these moments usually represent themselves through a perceptible change in the non-verbal domain.

Stern et al. (1998) proposed that there is "something more" than the classical mutative function of interpretation. They suggested the idea of "now moments" happening in the process of psychotherapy. These now moments usually have a feeling that is out of ordinary and create anxiety in therapist and patient since it often necessitates a choice to be made. Stern et al. state that "these moments are pregnant with an unknown future that can feel like an impasse or an opportunity" (p. 912). If attended, now moments become a "moment of meeting" and have the potential to alter the *implicit relational knowing* of the relationship.

Knoblauch (2000) focuses on the "musical edge" of the therapeutic relationship and in his book gives extensive illustrations of rhythms and melodies of his patient's communications. Knoblauch describes a therapeutic stance in which he closely attends to the subtle dynamics of his patients' speech and, just like a jazz musician would do, improvises along the incoming material. His introduction of musical metaphors such as dissonance, harmony and accompaniment, provides a terminology thorough which the non-verbal interaction between the therapist and patient could be described.

1.2. EMPIRICAL

1.2.1. Research on Temporal Characteristics of Speech

1.2.1.1. Early Period

The first research on the temporal characteristics of speech and dialogue was done by Norwine and Murphy (1938, as cited in Holtz, 2003) where they investigated telephone conversations. Their observations revealed that dyadic conversations have a turn-based structure; one party talked for a time, then the other party started talking usually following a pause. They gave different names to these events. The phase where each partner was talking was called a *talkspurt*. The pauses within talkspurts are called *resumption times* and the pause between one speaker and the other is called a *response time*. There were some instances where both speakers engaged in simultaneous talkspurts, which they named *double talking*. Theirs was the first attempt to operationalize the various events happening in spoken conversations, and following researchers have employed slightly modified versions of these definitions.

Next, Chapple (1949, as cited in Holtz, 2003) devised a system called the Interaction Chronograph, where the conversation of two people is monitored and recorded by an outside observer, who marked the beginnings of their vocal behaviors by pushing several buttons. Chapple and his team applied this methodology to study conversations between supervisor-supervisee, personnel interviewer-job applicant, and doctor-patient dyads. They found that individuals have characteristic rates of speaking and being silent and some people were more willing to change their own rates according to the other speaker.

Matarazzo et al. (1958) did a factor analysis of 12 variables of the Interaction Chronograph method devised by Chapple (1949, as cited in Holtz, 2003) and found that two factors were very stable for any individual: how long a speaker waits before communicating, and the number and average duration of communicative interactions.

In their studies, Matarazzo et al. (1958) used a human rater for the measurement of the variables like the previous group of researchers. In addition, they devised what they called "the standardized interview" method, where confederate interviewers are trained to use fixed durations of vocalizations or pauses during a particular period of the interview. In one of their experiments, they designed a 20-minute interview where the interviewer deliberately elongated the duration of his utterances in the middle 10-minute segment. They found that during this 10-minute interval, the interviewee also elongated the duration of his utterances (Matarazzo et al. 1968, as cited in Holtz, 2003). Although Chapple's (1949, as cited in Holtz, 2003) and Matarazzo et al.'s studies aren't specifically aimed at analyzing the moment-to-moment influence of one speaker on the other, their results showed that individuals have a tendency to match the duration of vocalizations with their partners in conversation (Matarazzo, Hess, & Saslow, 1962).

Matarazzo et al. (1968, as cited in Holtz 2003) also studied the interaction between the patient-therapist dyad within a psychotherapy context. They analyzed audio recordings from seven psychotherapy processes and measured several temporal variables including reaction time latency, which is the term they used to indicate the duration of silence between speaking turns; the duration of utterances, which is the duration of speech until a major pause; and incidence of interruption, which referred to the occurrence of simultaneous speech. With regard to the reaction time latency and incidence of interruption, they found a positive correlation between the patient and therapist in most of the cases. However, the durations of utterances were only correlated in one of the seven cases, and the direction of influence was negative. According to Holtz (2003), this was the first research where the temporal characteristics of speech are measured in a psychotherapy context.

The early period of research is characterized by a lack of a standardized definition of variables and the employment of human raters for their measurement. Therefore, Holtz (2003) argues that due to these problems it is difficult to compare the results of early studies. However, they were influential in their attempts in defining the variables, testing their stability over time and investigating interpersonal influence.

1.2.1.2. Rhythms of Dialogue

One breakthrough moment on the history of the research on temporal characteristics of conversation happened when Jaffe and Feldstein (1970) published *Rhythms of Dialogue*. They have developed a computerized method for the measurement of vocal timing, which they have called The Automatic Vocal Transaction Analyzer (AVTA). As a result of being able to use an automatic rating system instead of human observers, they were able to propose a standardized classification of events which are characteristic of every human vocal dialogue. First, they defined a conversation as "a sequence of sounds and silences generated by two (or more) interacting speakers" (p. 19). Those sounds and silences might be both unilateral (a speaker talking while the other one is listening) and bilateral (both speakers talking or silent at the same time). Secondly, a speaker is said to be gaining the "possession of floor" with his first unilateral utterance and holds it until the first unilateral utterance of the next speaker, who then gains the possession of the floor. Next, a computer program is used to classify the temporal events occurring in a dialogue into five categories: *vocalization*, *pause*, *speaker switch*, *switching pause* and *simultaneous speech*. A *vocalization* is defined as "a continuous sound by the speaker who has the floor" whereas a *pause* is "a period of joint silence bounded by the vocalizations of the speaker who has the floor". A *speaker switch* occurs when a speaker loses the possession of floor and the *switching pause* is a "period of joint silence bounded by the vocalizations of different speakers", which is assigned to the speaker who loses the floor. Lastly, a *simultaneous speech* is "a sound by the speaker who does not have the floor during vocalization by the speaker who does" (Jaffe & Feldstein, 1970, p. 19).

Jaffe and Feldstein (1970) tested the reliability of these parameters in various experiments and concluded that *vocalization, pause, speaker switch, switching pause and simultaneous speech* are distinct events and individuals have characteristic styles with regard to the duration of these five variables. Moreover, the level of stress, loss of visual cues, and change of conversational partners resulted in a decrease in stability of those parameters. Finally, their findings showed that, although individuals tend to match their average duration of pauses and switching

pauses with their partners, they did not match the duration of their vocalizations. They termed this influence of one speaker over the other as “congruence”.

In one of the experiments, Jaffe and Feldstein (1970) investigated the relationship between congruence of temporal patterning and individuals’ affinity for each other. In an experiment where participants conversed with each other for eight consecutive days, Jaffe and Feldstein observed that over time, participants started matching their duration of pauses. In addition, pairs that showed congruence rated each other as “warm, similar, and as someone they would invite to dinner” (p. 46). Similarly, Welkowitz and Kuc (1973) measured subjects’ congruence of the duration of vocalizations, pauses, and switching pauses in an interview context, by using AVTA. In addition, outside observers rated the conversations in terms of each subject’s warmth, empathy and genuineness. Their findings revealed a correlation between the congruence of switching pause and the level of warmth. The more individuals matched the duration between the speaking turns, the more they were rated as being warm by outside observers.

The index of interspeaker influence in the studies mentioned above is calculated by comparing the mean durations of events across the entire conversation; such as the mean duration of vocalization of a 10 or 45-minute period. This method cannot account for how individuals influence each other during the course of a specific time period on a moment-to-moment basis. In the next section, we will review some of the studies that focused on this aspect.

1.2.1.3. Research on the Moment-to-Moment Coordination of Rhythm

Cappella and Planalp’s (1981) study might be the first one to provide evidence for the existence of interspeaker influence on a moment-to-moment basis. Across 12 dyadic conversations they have measured the duration of vocalizations, pauses and switching pauses. They used two different methods to summarize the data. The first was the probability method, where they divided a 20-minute conversation into 2-minute blocks and calculated the probability of the occurrence of temporal events for each 2-minute period. The second was the turn method, where they calculated the duration of events in each speaking turn. They employed

time series regression analyses to calculate first; the effect of each individual's past probabilities on its current probability (the consistency effect) and secondly; the effect of one speaker's past probabilities on the other speaker's current probability (the influence effect). Their analyses showed the existence of interspeaker influence regardless of the method of summary being used. The magnitude of influence varied from dyad to dyad, and overall, they observed a "weak but detectable" level of influence (p. 127). Moreover, the direction of influence could be either positive or negative; positive coefficients meant "matching", while negative coefficients indicated "compensating". The most unequivocal conclusion that they arrived was that partners in a conversation was matching the duration of switching pauses on a moment-to-moment basis. They also found, to some degree, the matching of the duration of pause and compensation of the duration of vocalization between partners.

In order to examine the relationship between moment-to-moment coordination of temporal vocal parameters and outside observers' ratings of pleasantness of affect and the degree of involvement; Warner et al. (1987) measured the percent of time spent on vocalization in 15-second intervals during the course of 40-minute conversations. Time-series analysis revealed a curvilinear relationship between matching of the vocal rhythms and outside observers' rating the conversations as being "positive". In other words, conversations with moderate amounts of rhythmic matching were observed as being positive, whereas conversations with no or extreme rhythmic matching were evaluated as being negative. These findings supported the idea that distressed couples were exhibiting interaction patterns that are rigid and predictable (Gottman, 1979, as cited in Warner, 1987). This curvilinear relationship is also observed by some of the later researchers (Holtz, 2003; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001) and the term "optimum midrange" is used to refer moderate degrees of coordination between the partners. For an extensive review of this phenomenon see Beebe and McCrorie (2010).

A number of studies have investigated the coordination of temporal parameters between mothers and infants. Beebe et al. (2005) argue that the infant-adult and adult-adult dialogues are similar in terms of their overall temporal

structure and the way it is coactively constructed by the participation of its both members. Feldstein et al. (1993) have observed plays of 28 four-month-old infants with their mothers and strangers at two different sites. With the application of AVTA to the audio recordings of interactions, they have measured the duration of vocalizations, pauses and switching pauses. The durations of these variables were averaged for 5s intervals to create time series. Through regression analyses, they found significant degrees of coordination between mothers and infants, regardless of the sites. In addition, they found that a 20 to 30-second lag was best for predicting one partner's behavior from the other's. Similarly, Jaffe et al. (2001) assessed the coordinated interpersonal timing (CIT) of 88 four-month-old infants with their mothers and strangers in face to face interactions at home and lab settings. They also measured infants' attachment style with the Ainsworth Strange Situation when they are 12-months-old. Degree of rhythmic coordination at 4-months between infants and mothers predicted infant's attachment style at 12 months. Moreover, midrange degrees of coordination were found to be associated with the secure attachment type.

Holtz (2003) applied a similar methodology to study the self and interactive regulation of vocal rhythms in an adult-adult brief psychodynamic therapy setting. Holtz analyzed the audio recordings from three different psychotherapy processes through a slightly modified version of AVTA, which she called the Vocal Interaction Digital Audio (VIDA) system. For each session, duration of vocalization, pause, switching pause and the number of incidents of non-interruptive (NSS) and interruptive (ISS) simultaneous speech are measured. Holtz also introduced a new variable called the zero response latency (ZRL) events, which indicated the occurrence of turn switches without a switching pause. Overall, therapists' duration of turns and vocalizations are found to be much shorter than of the patients', indicating that the structure of psychotherapeutic dialogue was organized differently compared to regular conversations. She also found coupling (ordinary correlation) of NSS, ISS and ZRL events in two cases. In the third case, she found coupling of all of the variables. With respect to moment-to-moment vocal rhythm coordination; Holtz analyzed for each member of the three dyads, the degree of self-regulation (autocorrelation) and interactive regulation (cross correlation) of

vocalization, pause and switching pause. She found in all cases a matching of the switching pause, and a compensation of the duration of vocalization. The duration of pause was both matched and compensated. Holtz also looked for the relationship between the self and interactive regulation of these variables with measures of therapists' interpretive accuracy and patients' level of experiencing. Her findings tentatively supported the "optimum midrange" hypothesis, where midrange degrees of self-regulation of switching pause were associated with patients' therapeutic progress.

1.2.2. Research on Intensity and Pitch

1.2.2.1. Research on the "Quality of Voice"

We will now review several studies that employed human observers in the measurement of a construct that is usually called "the voice quality". The process of rating "voice quality" depends on outside observers' subjective perception of a combination of vocal parameters, but it is mainly influenced by the fluctuations of intensity and pitch. Since intensity and pitch contours play a significant role in the transmission of emotions (Scherer, 2003), it is no surprise that categories of voice quality are often named after the emotions they evoke.

Milmoe et al. (1967) asked doctors working at an Alcoholic Clinic how they felt about their patients. They recorded doctors' responses on tape and content-filtered it, which is a process where some frequencies of audio material are removed so that only the variations in intensity and pitch are discernable. They asked judges to rate those recordings for its emotional content. They found that doctors whose voices are judged with an anxious quality were more successful in referring alcoholics for further treatment than doctors with an angry voice quality.

Duncan et al. (1968) investigated voice qualities of therapists' responses from sessions that are rated either as being a peak hour or a poor hour by the therapists. Observers were trained to attend variations in intensity (overloud, normal or oversoft intensity), pitch height (overhigh, normal or overlow pitch) and

vocal lip control (normal or open lip control). They have found that vocal qualities associated with peak hours were open voice, oversoft intensity and overlow pitch.

Rice and Wagstaff (1967) and Rice and Kerr (1986, as cited in Tomicic, Martinez, & Krause, 2015) developed measures for assessing voice qualities of both clients and therapists. Rice and Wagstaff (1967) studied a wide range of client responses focusing on their intensity, pitch range, tempo, and stress patterns and identified four categories of client voice qualities (CVQs). The *emotional* category consisted of responses that had "energy overflow rather than control, (where) the voice breaks, trembles, or chokes". *Focused* responses had "a good deal of energy, but not a wide pitch fluctuation". Responses that had "high energy and a wide pitch range" were categorized as *externalizing*. And finally, responses with "low energy, a narrow pitch range and an even tempo" constituted the *limited* category (p. 558). Rice and Kerr (1986, as cited in Tomicic, Martinez, & Krause, 2015), applied a similar methodology to classify therapist voice qualities (TVQs). TVQ categories are: *softened*, "characterized by a lax voice that creates an intimacy and involvement effect"; *irregular*, "irregular intensity stresses with some pitch variation"; *natural*, "neither overly tense nor relaxed"; *definite*, "full, measured, assured, generally on the speaker's pitch platform"; *restricted*, "strained, slightly tremulous, whiny, droning…"; *patterned*, "patterned for emphasis, specially using pitch"; *limited*, "low energy… and a monotone pitch" (Tomicic, Martinez, & Krause, 2015, p. 3-4). Wiseman and Rice (1989) studied the interaction of TVQs with CVQs in change events. In their definition, a change event started at a point when the patient considered his or her reaction to a particular situation as problematic. They found that a therapist response with an *irregular* TVQ could predict a *focused* VCQ on the patient's part in change events.

Bernholtz (2013), applied CVQ and TVQ measures to selected episodes of cognitive behavioral therapy (CBT) and process-experiential emotion focused therapy (PE-EFT). Results indicated that patients CVQs had mostly *externalizing* quality in CBT, and an *emotional* quality in PE-EFT sessions. Overall, a combination of *emotional* and *focused* VCQs predicted better outcome results.

Those studies on "voice quality" have relied on human observers for the measurement of intensity and pitch. This approach lacks the precision that one finds

in the automatized measurement of temporal characteristics (Jaffe & Feldstein, 1970).

1.2.2.2. Automatized Measurement of Intensity and Intensity Convergence

The Automatic Vocal Transaction Analyzer (AVTA) developed by (Jaffe & Feldstein, 1970) scans the audio recordings at 250ms intervals and detects whether there is a sound or silence according to a predefined threshold value. For example, setting a 25 decibel (dBs) threshold meant that sounds that have amplitude levels below 25 dB were regarded as silences. A human operator set the threshold levels for each individual recording. While this method enabled filtering the desired sounds from other background noises, it produced a binary sequence of 1's and 0's representing sounds and silences. The fluctuations of intensity occurring in speech were not represented by this binary sequence.

To explore this, Welkowitz et al. (1972) designed an experiment where they measured the level of intensity of speakers who met once a week for three weeks. They have found that the intensity levels of interacting speakers remained fairly stable within a conversation. In this respect, they concluded that intensity was similar to the temporal characteristics of speech defined by Jaffe and Feldstein (1970). In addition, they found evidence that indicated the existence of intensity convergence. In other words, a speaker's level of intensity of vocalizations was matched by the other speaker. Up until that time, most studies on vocal congruence indicated that conversational partners were matching the duration of pauses, but not the duration of vocalizations. Welkowitz et al. propose that, as an alternative to the duration of vocalizations, the fluctuations in intensity could be used as an index of interspeaker influence.

Natale (1975) studied the relationship between intensity convergence and social desirability. Partners that are unknown to each other engaged in hour long unstructured conversations. Prior to the first conversation, they rated each other on a desirability scale. Results indicated that lowering or raising the level of intensity by one participant resulted in the other party to change the intensity of vocalizations

in the same direction. In addition, subjects who liked each other showed higher rates of intensity convergence.

1.2.2.3. Research on Pitch Synchrony

Reich et al. (2014) conducted the first research that investigated the relationship between pitch synchrony and outcome measures in a psychotherapy context. They first segmented the audio recordings of sessions into speaking turns and then measured the mean pitch level of each turn with the use of PRAAT, a computer software developed for speech analysis (Boersma & Weenink, 2019). They looked for both therapist-leading and therapist-following synchrony. For therapist-leading synchrony, they calculated the correlation between the mean pitch of a therapist turn and the mean pitch of the following patient turn, whereas for therapist-following synchrony they looked for the correlation between the mean pitch of a patient turn and the following therapist turn. They have found a modest degree of synchrony for both types, where the therapist-leading correlation was slightly higher. When they investigated the relationship between pitch synchrony and therapeutic alliance and outcome measures, they found that synchrony was negatively correlated with therapeutic alliance. The results were puzzling since it contradicted with their expectations that higher rates of non-verbal synchrony will be associated with better relationship quality. They argued that patients might have started to imitate the pitch of their therapist when they felt that the relationship is suffering. They suggested that future research should examine the moments that have high and low pitch synchrony in detail.

Imel et al. (2014) investigated pitch synchrony in role played psychotherapy sessions. They looked for the correlation between pitch levels of therapist and patient. Pitch levels are measured again by PRAAT, for every minute of the conversation. In addition, they compared the correlation coefficients with ratings of empathy. They found pitch synchrony between the patient and therapist and they observed higher levels of synchrony at sessions with higher empathy ratings. Moreover, they stated that sessions that are rated higher in empathy were characterized by lower levels of pitch in general.

1.2.2.4. Communicative Musicality

Malloch (1999) proposed the term "communicative musicality" referring to the musical qualities of the non-verbal interaction between mothers and infants. Their interaction is examined in terms of their pulse, quality and narrative which, for Malloch, are the "attributes of human communication, which are particularly exploited in music, that allow co-ordinated companionship to arise" (p. 32). Pulse and quality are measured by looking at spectrographs[3] extracted from the audio recordings of mother-infant interactions, that usually lasted one or two minutes. With regard to the interaction in the domain of pitch, close up analyses showed that the mother is aware of infant's pitch contour. Mothers followed infants' direction of movement; when the infant goes upwards in pitch, the mother exaggerated this with a greater curve. When it goes down, again, the mother followed the infant. In addition, pitch plots revealed recurring waves that lasted 20-25 seconds. Overall, pitch goes up and returns to the same spot in the span of this 20-25 seconds. Malloch argued that this was an indication that partners are aware of both micro and macro rhythms.

1.3. SUMMARY OF LITERATURE AND EXPLORATORY HYPOTHESES

Following Freud's exclusion of the sonic dimension from the psychoanalytic practice, a number of writers have advocated the consideration of non-verbal properties of speech for various purposes. Their purposes can be roughly categorized into three groups: 1) As a tool for the diagnosis of psychopathology (Moses, 1954, as cited in Bernholtz, 2013). 2) As a way of understanding the dynamics of patients' internal world where their verbal communications appear as insufficient or misleading (Reik, 1949, as cited in Templeman, 1977; Maiello, 1995). 3) As a method of therapeutic action, relying on the mutative effects of the implicit dimension of relationship that is co-constructed by the therapist and patient

[3] A spectrograph is a tool that visualizes the progress of amplitude and frequency of sound. In other words, it shows the movement of intensity and pitch in time.

(Beebe & Lachmann, 2002; Boston Change Process Study Group et al., 2002; Knoblauch, 2000).

Empirical research conducted on this topic has an early period where the measurement of vocal qualities, such as rhythm, intensity or pitch, relied on ratings of human observers (Duncan et al., 1968; Matarazzo et al., 1962; Milmoe et al., 1967; Rice & Wagstaff, 1967). The lack of standardized operationalization of variables make it difficult to compare the results and draw conclusions from this period of research. With the introduction of computerized measurement of vocal parameters and employment of time-series analyses researchers began to analyze the moment-to-moment interactions between the patient and therapist, or mother and infant (Feldstein et al., 1993; Holtz, 2003; Imel et al., 2014; Jaffe et al., 2001; Jaffe & Feldstein, 1970; Reich et al., 2014; Welkowitz et al., 1972; Welkowitz & Kuc, 1973). Results of these studies indicate that, in general, patient and therapist tend to match the duration of pauses and switching pauses, as well as the level of pitch and intensity. On the other hand, the durations of vocalizations are usually compensated, meaning that they are negatively correlated. A phenomenon that researchers usually stumbled upon in the research of the coordination of vocal parameters is summarized with the concept of "the optimum midrange" (Beebe & McCrorie, 2010). Higher levels of coordination or synchrony in rhythm and pitch is often found to be an indication of a troubled relationship (Holtz, 2003; Jaffe et al., 2001; Reich et al., 2014).

It should be noted that the number of studies that focused on temporal characteristics of speech, such as the duration of vocalizations and silences, are much larger than the studies that focused on intensity or pitch. Similarly, the number of studies that have been done using the recordings of actual psychotherapy sessions are less compared to studies done in experimental interview contexts. Moreover, researchers often focused on a particular segment of a session, as opposed to the consideration of the whole session (except Holtz, 2003).

In this study we used audio recordings of psychoanalytic psychotherapy sessions done with a 17-year adolescent for the measurement of the variations in rhythm, intensity and pitch over the course of 45-minute sessions. This is an exploratory study; therefore, no hypotheses were formulated prior to analyses.

The first aim of this study is to provide an overall description of a psychoanalytic psychotherapy process in terms of its rhythm, intensity and pitch. Six sessions from early, middle and late phases of treatment will be selected. For each session and for each member of the dyad, the number and average duration of temporal events such as turns, vocalizations, silences, switching pauses; as well as mean level, range and standard deviation of intensity and pitch of the session will be presented. This first goal of the study is an attempt to answer following questions such as: How is the temporal organization of these sessions? How long, on average, does each member of the dyad possess the speaking floor and how long they wait before switching turns? What is their average duration of vocalization and pauses in a turn? Do these temporal variables, alongside the mean level of patient's and therapist's intensity and pitch, differ between early, middle, and late periods treatment?

Our second objective is to depict how some of these parameters change within the course of each session. Through the graphic illustration of moment to moment variations in intensity and pitch, our aim is to observe individual trends within a session. In addition, we want to assess whether an individual's vocal behavior at a given time with regard to its rhythm, intensity and pitch can be predicted from its previous states.

Our third goal is to see how much a member of the dyad's vocal behavior is correlated with the other's behavior during the course of a session. Do the therapist's variations in rhythm, intensity and pitch fluctuate in relation to the patient's, or vice versa? In other words, can a member of the dyad, or both, be thought of as "following and responding to the other"? If they are "following", do we see a tendency to match or is there a divergence? Moreover, we want to observe how this "matching" behavior varies between the early, middle and late periods of therapy.

In the fourth and final aim of this study we will first identify moments where the patient's and therapist's intensity and pitch are at its minimum and maximum level. After going back to the original recordings and examining the verbal interaction that is happening at those peak moments, we will present a transcript of one selected segment for each session. The qualitative analysis of the verbal

material will be used to interpret the dynamics happening at the non-verbal domain, or vice versa.

CHAPTER 2
METHOD

2.1. SAMPLE

In this study, we used audio recordings and transcripts of sessions that have been done with a 17-year-old female adolescent. The data was collected for the Psychotherapy Research Program in Istanbul Bilgi University. Participants' consents are obtained prior to beginning of the treatment. Ms. H. applied to the clinic with depressive symptoms and an unwillingness in attending the school. She was diagnosed by the staff psychiatrist as having Major Depressive Disorder. The psychiatrist also prescribed her anti-depressants. The psychotherapist that worked with Ms. H was a member of the clinic, who had 4 years of previous experience. The theoretical orientation of treatment was psychoanalytic psychotherapy. Ms. H. regularly attended once-a-week sessions for 11 months. The treatment was prematurely terminated because the therapist had to move to another country for educational purposes.

6 sessions are selected from a total of 43 available sessions, representing early, middle and late phases of treatment. Session 1 was not recorded since it was the intake session. Sessions 2 and 4 were sessions that are conducted with the mother of the patient. Sessions 3 and 5 were recorded but audio files were damaged. Session 7 was not recorded due to a technical error. As a result, sessions 6 and 8 were the first sessions that were available; therefore, these two sessions are selected in order to represent the early phase. Next, sessions 45 and 46 are selected from the late phase. Finally, sessions 26 and 27 were selected from the middle phase.

2.2. MEASURES

PRAAT (Boersma & Weenink, 2012), a computer software developed especially for speech analysis was used in order to measure the variables related to rhythm, intensity and pitch. A PRAAT script has been written for this study for the automatization of measurement process.

Under the category of rhythm, several temporal variables are defined and measured. Those variables are; *turn, vocalization, pause*, and *switching pause*. The operationalization of these variables is influenced by Jaffe and Feldstein (1970), but they are not identical. Due to technical impossibilities, we did not use variables that are related to simultaneous speech, such as interruptive simultaneous speech or non-interruptive simultaneous speech.

A speaking *turn* starts with the first word that is uttered by an individual and ends with the first word uttered by the other. Sounds that are not words, such as "hm"s or "uh-huh"s do not start or end a turn. Any type of sound produced by the turn holder having an intensity higher than a specific threshold value is counted as a *vocalization.* The threshold value is set individually for each session, depending on the level of background noise. Silences that are bounded between two vocalizations are coded as a *pause.* The silence at the end of a speaking turn is coded as a *switching pause*. The *switching pause* is a time period that is coordinated mutually by the two participants; one starting it by finishing her talk, and the other ends it by beginning speaking. If the turn ends without a switching pause, this indicates that either a speaker started vocalizing without waiting the other to finish, or the switch happened immediately after one speaker finished her vocalization.

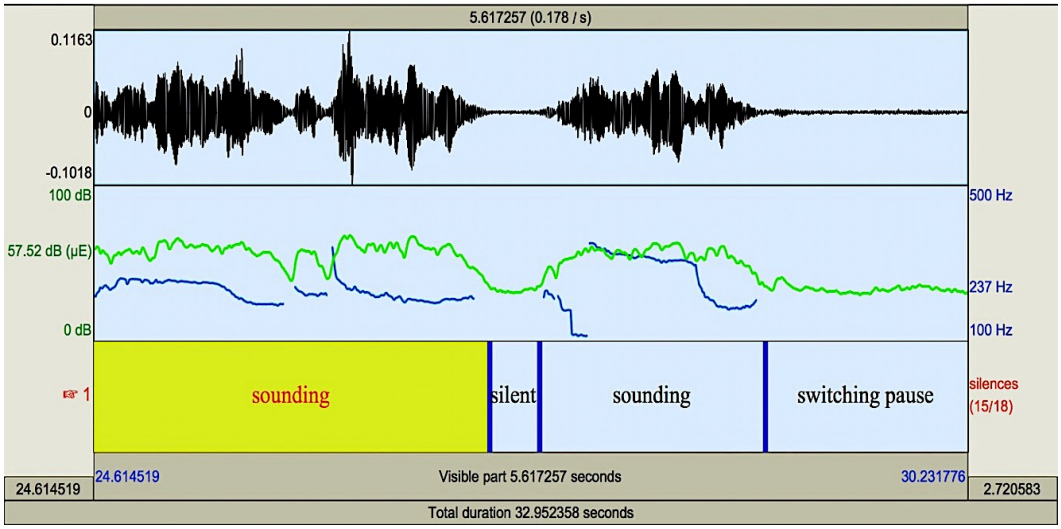
The minimum duration of vocalization, silence and switching pause is 250 milliseconds. This duration is used because previous studies used similar intervals (Holtz, 2003; Jaffe et al., 2001; Jaffe & Feldstein, 1970). If a silence between vocalizations is less than 250ms's, it is not coded as a pause. Instead, that silence is considered as a part of the vocalization period. Similarly, a sound shorter than 250ms's does not count as a vocalization.

A *length* of vocalization, silence or switching pause indicates the duration of vocalization, silence or switching pause without any interruption. For example,

a length of 2.5 seconds of vocalization means that there are no pauses present during 2.5 seconds of continuous vocalization. On the other hand, when we refer to a *duration* of vocalization, or pause in a speaking turn, it is the total duration of all vocalizations or pauses within a turn.

Intensity is measured by from the amplitude level and pitch is measured from the fundamental frequency (F0) of the signal. Figure 2.1. shows the waveform, pitch and intensity curves, and the classification of the waveform into sounds and silences.

Figure 2.1. Waveform; intensity and pitch curves; and the categorization into sound, silence and switching pause.



2.3. PROCEDURE

Since the initial audio recordings have been done by using a single recording device that is placed between the speakers, voices of both members are inevitably recorded on a single track. Therefore, first a process of segmentation was applied where the patient and therapist's voices are separated into two different audio tracks. Any type of sound uttered by speakers (including back-channel responses such as "uh-huh"s) is separated into two tracks. In addition, the beginning and endings of each turn is marked so that we had each member's sounds for each turn. The process of segmentation took approximately 2 to 3 hours for each session.

After the segmentation, we used two different methods for the analysis of the audio material. First, in the turn-based method, the PRAAT script calculated for each turn the following variables: the duration of the turn, the number and total duration of vocalization and pauses within the turn, the duration of switching pause if there is any; mean, standard deviation and range of intensity; mean, standard deviation and range of pitch. Secondly, in the time-based method, the script analyzed the whole uninterrupted audio file of the session for the patient and therapist. In this method, the mean intensity and pitch for every 250ms is calculated. For a 45-minute session, the result is two series with approximately 10800 values.

2.4. DATA ANALYTIC PROCEDURES

In order to assess whether an individual's vocal behavior at a given time was predictable from his or her states at a previous moment we have calculated the autocorrelation of a given time-series. To assess whether the patient or therapist is following and matching his or her vocal behavior with each other, we have calculated the cross correlation between the time-series of the patient and the time-series of the therapist. Under the category of rhythm, we have created five different time-series for each member representing the duration of vocalization, duration of pause, length of vocalization, length of pause and length of switching pause. Since we have measured these variables for each speaking turn, the number of time points in the time-series is equal to the number of turns. In order to create time-series for intensity and pitch, we have calculated the average level of intensity and pitch for each 30-second segment of the session. As a result, our time-series for intensity and pitch are composed of approximately 80-90 time points for each 45-minute session. There were a number of 30-second segments where the patient or therapist did not produce any type of sound and these are coded as missing values. In order to assess the stationarity of each time-series, we have conducted a series of unit root tests. For each time-series, Augmented Dickey-Fuller Test rejected the null hypothesis ($p < .05$) thereby indicating that the series are stationary.

CHAPTER 3
RESULTS

3. 1. DESCRIPTIVE RESULTS ACROSS SIX SESSIONS

Our first aim was to globally describe a psychoanalytic psychotherapy process in terms of its organization of rhythmic events, intensity and pitch. Table 3.1. summarizes the rhythmic events.

With respect to the number of turns (M= 89.50, SD=21.30), sessions 6 and 26 stand above average, indicating a faster, more interactive turn taking rhythm in those sessions. Overall, the therapist (M=10.99, SD=3.62) spends much less time than the patient (M=49.30, SD=9.33) in the possession of turn floor as expected by the rhythm of a psychoanalytic therapy. Figure 3.1. shows for each session the average turn duration, average duration of vocalizations and average duration of silences (sum of pause and switching pause) within a turn.

Figure 3.1. Mean Turn Duration and Mean Duration of Vocalization and Silence Within a Turn

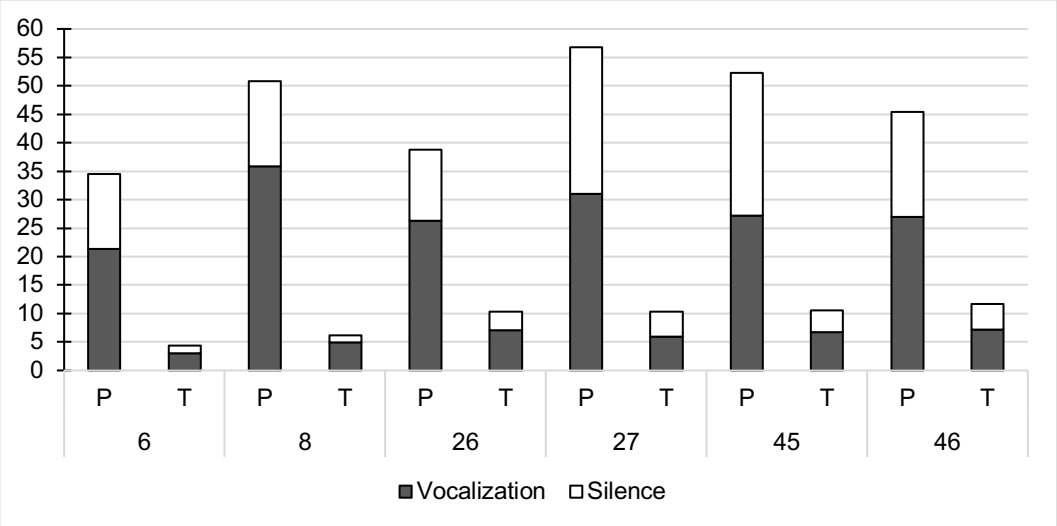


Table 3.1. Rhythm Variables Across Six Sessions

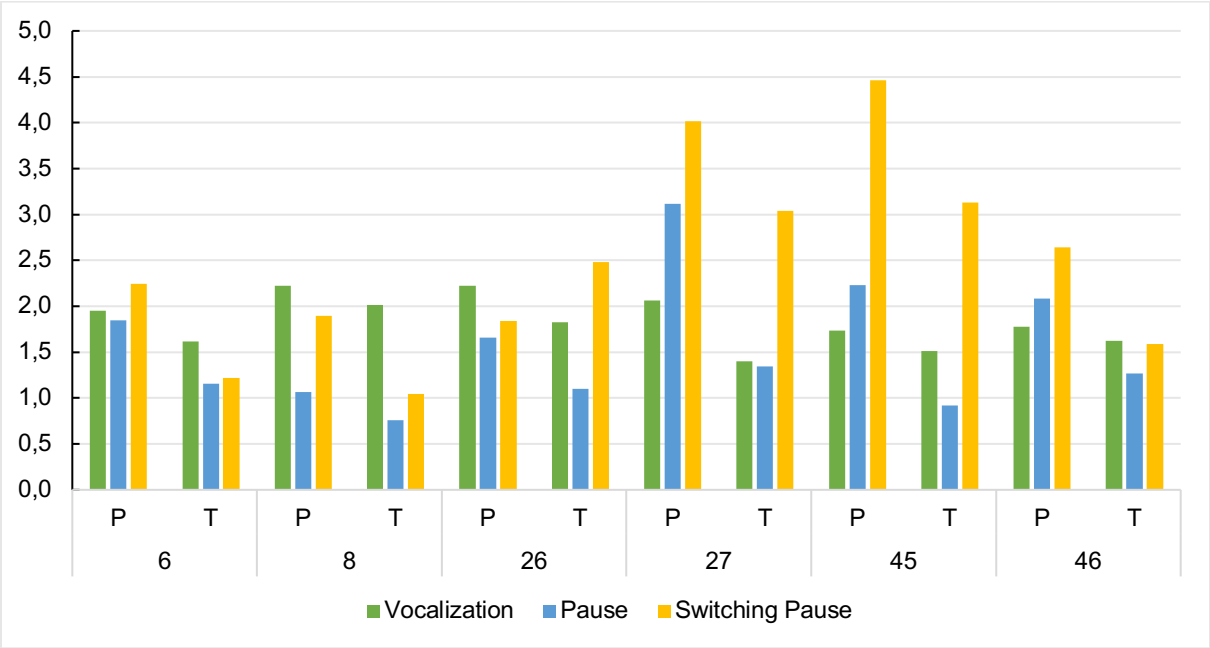
	6th Session		8th Session		26th Session		27th Session		45th Session		46th Session	
	P	T	P	T	P	T	P	T	P	T	P	T
Number of Turns	64	64	41	41	50	50	37	37	36	36	40	41
Mean Turn Dur.	36.73	5.58	52.70	7.19	40.64	12.84	60.85	13.39	56.78	13.69	48.10	13.27
Total Dur. of V	1367.04	195.36	1470.81	203.00	1313.78	353.34	1057.11	220.21	978.91	244.27	1079.03	288.02
Total Dur. of P	840.15	83.10	612.14	48.98	626.41	194.48	1045.29	162.42	904.40	135.73	786.91	179.26
Total Dur. of SP	143.61	78.04	77.81	42.88	92.03	124.06	149.04	112.52	160.68	112.73	105.59	63.59
Mean Dur. of V	21.36	3.05	35.87	4.95	26.28	7.07	31.04	5.95	27.19	6.79	26.98	7.20
Mean Dur. of P	13.13	1.30	14.93	1.19	12.53	3.29	30.72	4.39	25.12	3.77	19.67	4.48
Mean Length of V	1.95	1.62	2.23	2.02	2.22	1.83	2.07	1.40	1.73	1.51	1.78	1.62
Mean Length of P	1.84	1.15	1.06	0.76	1.66	1.10	3.11	1.35	2.23	0.92	2.08	1.27
Mean Length of SP	2.24	1.22	1.90	1.05	1.84	2.48	4.02	3.04	4.46	3.13	2.64	1.59

V = Vocalization, P = Pause, SP = Switching Pause
P = Patient, T = Therapist

Inspection of Figure 3.1. tells us that the therapist's average turn duration is getting longer as the therapy progresses. In addition, the patient's percentage of time spent vocalizing starts to become equal to the percentage of time spent being silent within a turn. A similar pattern is also observable in the therapist.

Figure 3.2. illustrates the mean lengths[4] of vocalization, pause and switching pause. Therapist's and patient's lengths of rhythmic events are much similar in comparison to their turn durations or total duration of vocalization and silences. But still, therapist's length of vocalizations and pauses are shorter in general. Except only in 26th session, therapist's length of switching pause is longer than the patient's. By averaging the values from the six sessions; we could tentatively propose that the patient's baseline rhythm of vocalization (V), pause (P) and switching pause (SP) is V: 1.99, P: 1.99 and SP: 2.85. On the other hand, the therapist's baseline rhythm is V: 1.66, P: 1.66 and SP: 2.08.

Figure 3.2. Mean Length of Vocalization, Pause and Switching Pause Across Six Sessions



In sessions 6, 8 and 26 the length of patient's vocalization is longer than its length of pause. It could be argued that this pattern is characteristic of a forward pushing beat, possibly a sign of a sense of urgency. Beginning with the 27th session we see a change in this pattern, where length of pause becomes longer than the length of vocalization. This pattern could

[4] Note that lengths represent the smallest continuous rhythmic events. For example, a length of vocalization shows the duration of vocalization uninterrupted by any other event. In this respect, they are similar to musical notes and rests.

indicate the emergence of a laid-back, “breathing” rhythm, which has more space between the utterances. However, the therapist does not match this beat. We will trace the roots of this change when we focus on session 27 in detail. Length of switching pause increases on sessions 27 and 45 for both the patient and therapist. In addition, it looks like there is a convergence on the durations of all three events. This convergence is especially visible on the switching pause, except 26th session.

Following the description of rhythmic events, we now move on to the parameter of intensity. Table 3.2. and Figure 3.3. summarizes the mean level, standard deviation and range of intensity of turns, across the six sessions.

Figure 3.3. Mean, Standard Deviation and Range of Intensity Across Six Sessions

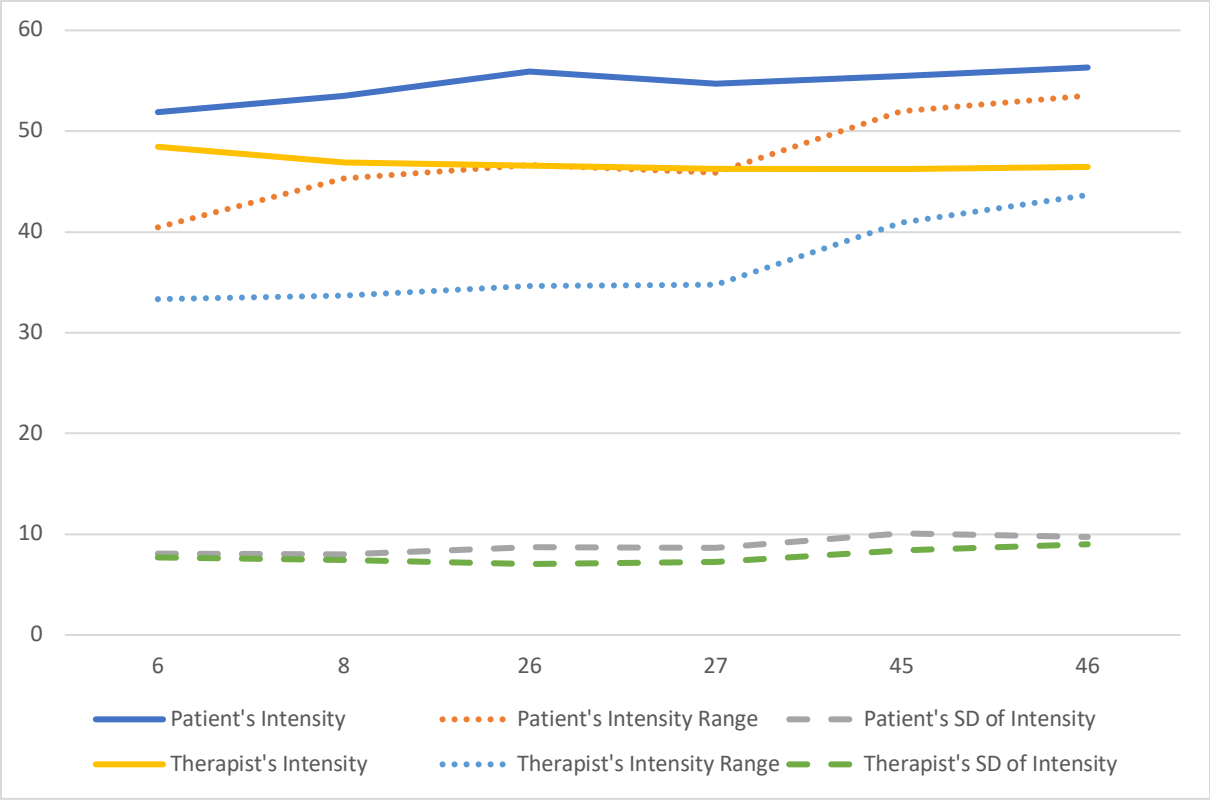


Table 3.2. Intensity Across Six Sessions

	6th Session		8th Session		26th Session		27th Session		45th Session		46th Session	
	P	T	P	T	P	T	P	T	P	T	P	T
M	51.89	48.45	53.51	46.88	55.93	46.55	54.72	46.26	55.50	46.24	56.32	46.44
SD	8.07	7.69	8.00	7.44	8.72	7.04	8.63	7.25	10.08	8.40	9.74	9.00
Range	40.45	33.33	45.29	33.68	46.63	34.65	45.90	34.78	51.99	40.91	53.52	43.66

P= Patient, T= Therapist

Table 3.3. Pitch Across Six Sessions

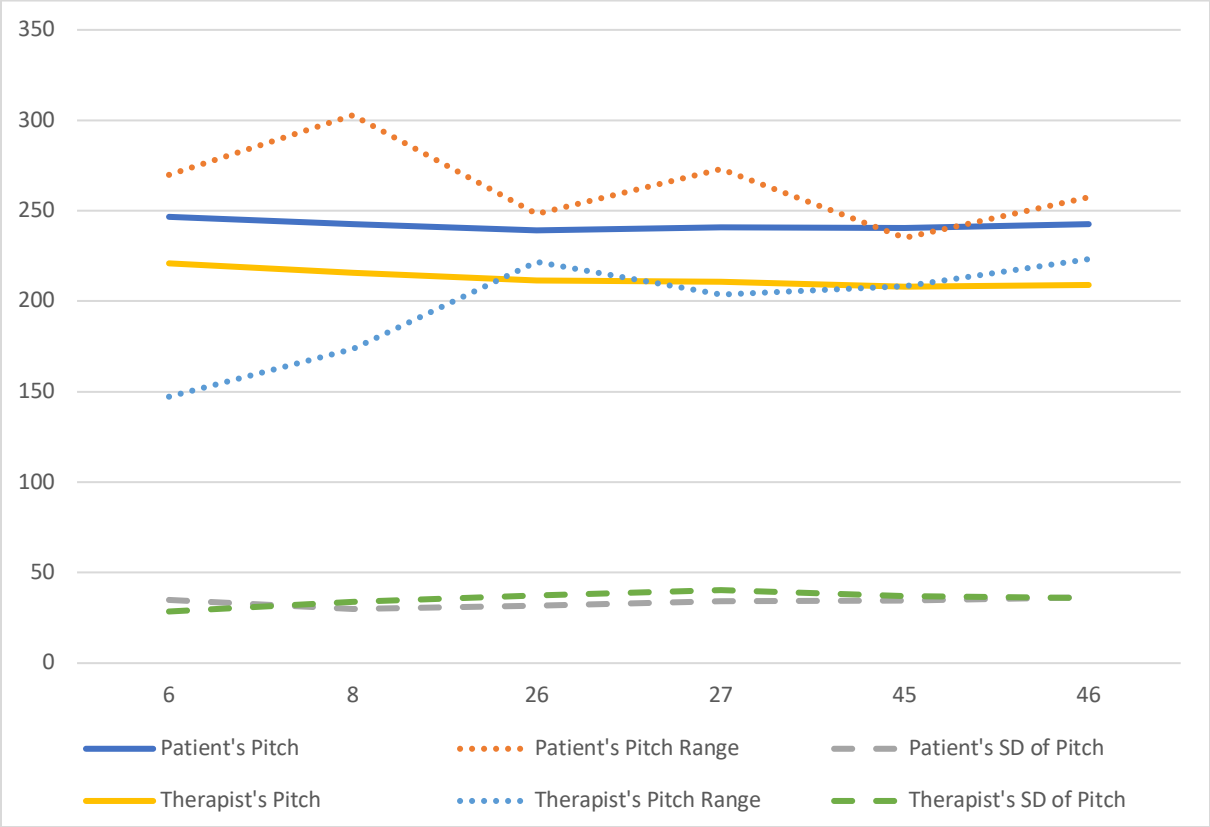
	6th Session		8th Session		26th Session		27th Session		45th Session		46th Session	
	P	T	P	T	P	T	P	T	P	T	P	T
M	246.66	220.95	242.60	215.64	239.19	211.33	240.79	210.91	240.60	208.06	242.68	209.15
SD	32.15	27.84	29.47	33.02	30.89	37.39	33.60	38.94	34.47	35.48	35.78	35.32
Range	269.91	147.26	302.83	173.50	248.30	221.79	272.96	203.60	235.08	208.15	257.61	223.30

P=Patient, T=Therapist

We see that the patient's level of intensity, standard deviation and range is higher than the therapist's. In addition, patient's level of intensity slightly increases as time progresses, whereas the therapist's level of intensity stays stable. In sessions 45 and 46, the range of intensity increases for both patient and therapist.

Next, we will move to the parameter of pitch. Table 3.3. and Figure 3.4. summarizes the mean level, standard deviation and range of pitch in turns, across the six sessions.

Figure 3.4. Mean, Standard Deviation and Range of Pitch Across Six Sessions



It seems that patient and therapist each possess a characteristic level of pitch that stays fairly stable over time. On the other hand, their pitch range varies from session to session. At the early phase of treatment, in sessions 6 and 8, the difference between the patient's pitch range and therapist's pitch range is at its highest. The therapist might have been using a constricted range of pitch due to various reasons that we'll further explore in the following sections. Starting in the middle phase of the treatment, their pitch range becomes much similar, still therapist's range being lower than the patient.

3.2. AUTOCORRELATION (SELF-REGULATION) OF RHYTHM, INTENSITY AND PITCH

Our second aim was to observe whether there is a self-regulation of rhythm, intensity and pitch. To examine self-regulation, we have analyzed the autocorrelation of the time-series. Under the category of rhythm, we looked for the autocorrelation of five variables: duration of vocalization, duration of pause, length of vocalization, length of pause and length of switching pause. For these variables, we will report the first statistically significant autocorrelation within two lags. Since for the rhythm variables we used a turn-based method, two lags span a time period up to four turns. For intensity and pitch, we will report autocorrelations up to three lags. Since for these two variables we used time-series that had 30 second increments, three lags span a time period of 1.5-minutes.

As for the autocorrelation of rhythm variables, it is hard to observe a consistent pattern that is present throughout the therapy process. For the therapist, we could not detect a statistically significant autocorrelation of rhythmic variables except 27th session. Session 27 is the only session in our sample, in which the therapist's rhythmic behavior could be predicted from its past behavior. In that session, we found lag (2) autocorrelations in three of the five rhythmic variables for the therapist. This suggests that in session 27, the therapist has moved out of her ordinary way of functioning due to reasons that we'll explore in further sections when we discuss that session in detail. It is interesting that we found autocorrelation for the therapist in this session for all of the "duration" variables and for one "length" variable, in which the direction of correlation was found to be negative. This observation suggests that in the exceptional case where the therapist's rhythmic behavior could be predicted from its earlier rhythmic behavior, her self-regulatory attempts are mainly related to the *duration* of vocalizations and pauses within a turn.

For the patient, we found lag (1) autocorrelation of length of pause in sessions 6 and 8, as well as lag (1) autocorrelation of length of switching pause in sessions 26 and 27. In addition, we found lag (2) autocorrelation of the duration of vocalization in session 26. As opposed to the therapist, patient's regulatory attempts seem to concern the *lengths* of pauses and switching pauses. As for the difference between the sessions, most autocorrelation in variables of rhythm come from the middle section of treatment and there is no autocorrelation in variables of rhythm in sessions from the late phase of the treatment.

Table 3.4. Autocorrelation of Rhythmic Events, Intensity and Pitch

	6th Session		8th Session		26th Session		27th Session		45th Session		46th Session	
	P	T	P	T	P	T	P	T	P	T	P	T
Dur. of V	0.05	0.17	0.03	0.05	0.29*	-0.07	-0.12	0.32*	-0.11	-0.03	-0.06	0.03
Dur. of P	0.00	0.01	0.02	-0.05	-0.16	-0.01	-0.06	0.51*	-0.19	0.05	-0.05	-0.07
Length of V	-0.11	-0.03	-0.14	-0.01	-0.15	0.08	-0.17	-0.30*	-0.10	-0.06	-0.13	0.19
Length of P	0.32*	0.04	0.40*	0.00	0.04	0.03	0.02	0.15	-0.08	-0.14	-0.01	0.15
Length of SP	-0.05	0.13	-0.14	0.05	0.32*	-0.09	0.37*	-0.06	-0.17	-0.11	0.09	-0.10
Intensity	0.34*	0.06	0.26*	-0.07	0.31*	-0.05	0.46*	0.03	0.19	-0.08	0.38*	-0.19
Pitch	0.70*	0.27*	0.17	-0.06	0.29*	0.12	0.26*	0.05	0.32*	0.07	-0.06	-0.01

V= Vocalization, P= Pause, SP = Switching Pause

P= Patient, T= Therapist

*$p < 0.05$

For the autocorrelation in the domains of pitch and intensity, it appears that there is a consistent pattern where a statistically significant lag (1) autocorrelation is present for the patient and not for the therapist. We observed three exceptions to this pattern. First, in session 6, the therapist also has lag (1) autocorrelation of pitch. Note that this is also the session where the level of autocorrelation of patient's pitch is at its highest ($r = 0.70$). Secondly, there is no autocorrelation of pitch for the patient, in session 8 and 46. Thirdly, no autocorrelation has been found for the level of intensity of the patient, in session 45.

Overall, almost all of the statistically significant autocorrelations that we found belong to the patient. This suggests that the patient's behavior in the domains of intensity and pitch is usually predictable from her behavior that is 30-seconds earlier. As for the rhythmic variables, it is hard to assume that such regularity exists. Nevertheless, it seems that the patient's rhythmic behavior is predictable with regard to the length of pauses, in the first phase, and length of switching pauses[5] in the middle phase. On the contrary, the therapist's vocal behavior is not usually predicted from its earlier behavior. There are some exceptions, and these will be addressed in the following sections.

3.3. CROSS CORRELATION (MUTUAL REGULATION) OF RHYTHM, INTENSITY AND PITCH

Our third aim was to observe whether the patient's vocal states with regard to its rhythm, intensity and pitch is predicted from the vocal behavior of the therapist, or vice versa.
We also wanted to see whether their behavior is of a matching type or a compensating type. We employed cross correlation analyses between the time-series of the patient and time-series of the therapist. For the variables that are related to rhythm we used turn-based time-series. Since the sessions we have sampled in this study always started with a patient turn, a lag (0) cross correlation indicates that the therapist is following the previous turn of the patient. On the other hand, a lag (-1) cross correlation means that the patient is following the previous therapist turn. In our analyses, we included the lags (-2), (-1), (0) and (1) for the rhythmic variables, spanning a period of time up to four turns. For intensity and pitch, since we used 30-second time-series, a lag (0) cross correlation means that the patient and therapist's behavior is correlated for the

[5] What we refer by a patient's switching pause is a type of pause that is initiated by the patient and collaboratively sustained by the therapist and patient by not engaging in another vocalization. By its nature, a switching pause is a phenomenon that is regulated by both participants. Therefore, an autocorrelation of the switching pause that we observed in the middle phase of treatment concerns both of the participants.

same 30-second period. We used lag (-2), lag (-1), lag (0), lag (1), lag (2) in our analysis, that encompassed a 1-minute period.

In all of the sessions, except session 26[6], we found cross correlation at lag (0) for at least one of the rhythmic variables. This suggests that throughout the psychotherapy process the therapist is consistently following the patient's rhythm of the previous turn, although the parameter on which this following occurs varies from session to session. We see a general tendency of matching the durations of rhythmic events, except only in session 8, where the length of vocalization is compensated. The therapist is observed as following and matching the duration of vocalization, duration of pause, length of pause and length of switching pause of the patient at various times throughout the psychotherapy process. Session 27 has the maximum number of variables (n = 3) in which the therapist is following and matching the patient's rhythmic behavior.

On the other hand, the cross correlations that are statistically significant in lags (-1) and (-2), indicating the patient following the therapist's rhythm, are much less in quantity. In the first half of the six sessions, we see no indication of patient following the therapist in terms of rhythm. We first observe a cross correlation in lag (-2) of the duration of pause in session 27, which is also present in the sessions after session 27. In other words, beginning in session 27 and continuing in the following sessions, the patient matches the duration of her pauses with the therapist's duration of pauses of two preceding turns. In addition, in session 46, a lag (-2) cross correlation of the length of vocalization, and a lag (-1) cross correlation of the length of switching pause are found. This indicates that alongside the duration of pause, the patient is following and matching the length of vocalization and length of switching pause of the therapist.

With regard to intensity and pitch, no cross correlation is found except two occasions. First, in session 8, a lag (-1) cross correlation of intensity; and secondly a lag (-2) cross correlation of pitch is found to be statistically significant ($p < .05$). In both of these situations, the patient is following the therapist with a matching type behavior.

[6] Note that in session 26 no cross correlation exist in any of the parameters.

Table 3.5. Cross Correlation of Rhythmic Events, Intensity and Pitch

	6th Session		8th Session		26th Session		27th Session		45th Session		46th Session	
	-lag	+lag	-lag	+lag	-lag	+lag	-lag	+lag	-lag	+lag	-lag	+lag
Dur. of V	0.02	0.14	-0.02	0.09	-0.09	-0.06	-0.20	0.09	0.08	0.56*	-0.19	-0.05
Dur. of P	0.02	-0.01	-0.08	0.40*	-0.04	-0.17	0.72*	0.41*	0.42*	0.15	0.37*	-0.14
Length of V	-0.18	0.22	0.07	-0.43*	0.00	0.17	0.31	-0.24	0.07	0.26	0.34*	-0.16
Length of P	0.07	0.34*	-0.21	0.07	0.24	-0.11	-0.04	0.52*	0.10	0.28	0.20	-0.14
Length of SP	0.01	0.12	-0.01	-0.06	0.07	0.16	0.28	0.41*	-0.18	-0.06	0.43*	0.35*
Intensity	-0.12	-0.11	0.25*	0.01	-0.06	0.06	0.07	0.17	0.03	-0.01	-0.20	0.19
Pitch	-0.19	-0.13	-0.06	-0.13	0.04	0.07	-0.13	-0.17	0.01	-0.18	0.26*	-0.12

V= Vocalization, P= Pause, SP = Switching Pause

+lag = Therapist following the patient

-lag = Patient following the therapist

*p < .05

An inspection of Table 3.5. suggests that, in terms of cross correlation, session 26 stands as a dividing point, in which there is no cross correlation of any of the variables. Sessions before 26 have less cross correlation of variables compared to sessions after 26. In a similar vein, the patient-following behavior seems to increase in the second half of the treatment, as the therapy process approaches to an end. On the other hand, the therapist seems to be following the patient on a regular basis, throughout the treatment.

3.4. SESSION SUMMARIES AND QUALITATIVE ANALYSES

In this section we will summarize the results that we have presented up to this point by focusing on each individual session. In each session we have attempted to identify the most dominant non-verbal theme of the session and interpret the results according to this theme. In this section we will also introduce graphs that illustrate the changes in pitch and intensity over the course of a session, and transcripts of verbal interactions from selected segments. The process of selecting relevant segments is carried out by first converting the time-series of intensity and pitch into z-scores and then identifying several maximum and minimum points. The time points that stand above 2 z-scores are selected as maximum points and time points that are below 2 z-scores are selected as minimum points. In cases where there are no time points that met these criteria, we have selected the top two maximum and top two minimum time points. Then, for each session, after investigating the content of the verbal conversation happening at these minimum and maximum points, we chose the ones that are thematically most relevant with the dominant non-verbal theme of the session.

3.4.1. Session 6

Non-verbal theme: “Fast turn rhythm and pitch regulation”

With a total of 128 turns and having the smallest average turn duration among the six sessions, session 6 is the one with the fastest turn taking tempo. This session is characterized by short and frequent bursts of speech that indicate the presence of a high amount of interaction between the partners. The length of pause variable seems to be the focus of rhythmic regulation. Other important characteristics of session 6 concerns the domain of pitch. In this session, we find the strongest autocorrelation of pitch in the patient ($r = .70$). This is also the only time where we see an autocorrelation of pitch in the therapist ($r = .27$). In addition, the difference between the pitch ranges of patient and therapist is one of the largest (122.65 Hz), where the therapist’s pitch range (147.26 Hz) is at its lowest among the six sessions. It could be argued

that the therapist is using a constricted pitch range in an attempt to regulate patient's heightened emotional arousal. In the first 3-minutes of this session, the patient's level of pitch increases gradually and then drops back to its normal levels around the 5th minute (see Figure 3.6.). In addition, we observe a gradual decrease in patient's pitch level over the course of the session. The therapist, on the other hand, stays fairly consistent in terms of its pitch level, with occasional rises and falls. Only in the last part of the session, the therapist seems to increase her pitch in a small amount. The non-existence of a cross correlation of pitch also suggest that the therapist does not get carried away by the emotional arousal of the patient and provides a stable frame for the patient to hold on to.

Figure 3.5. Intensity of Patient and Therapist in 6th Session

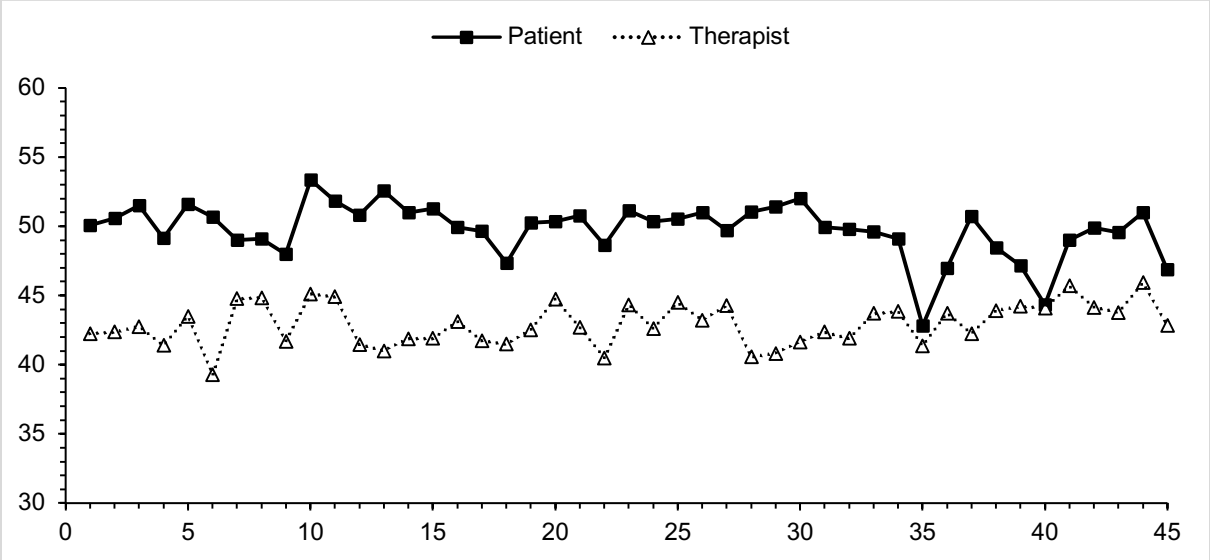
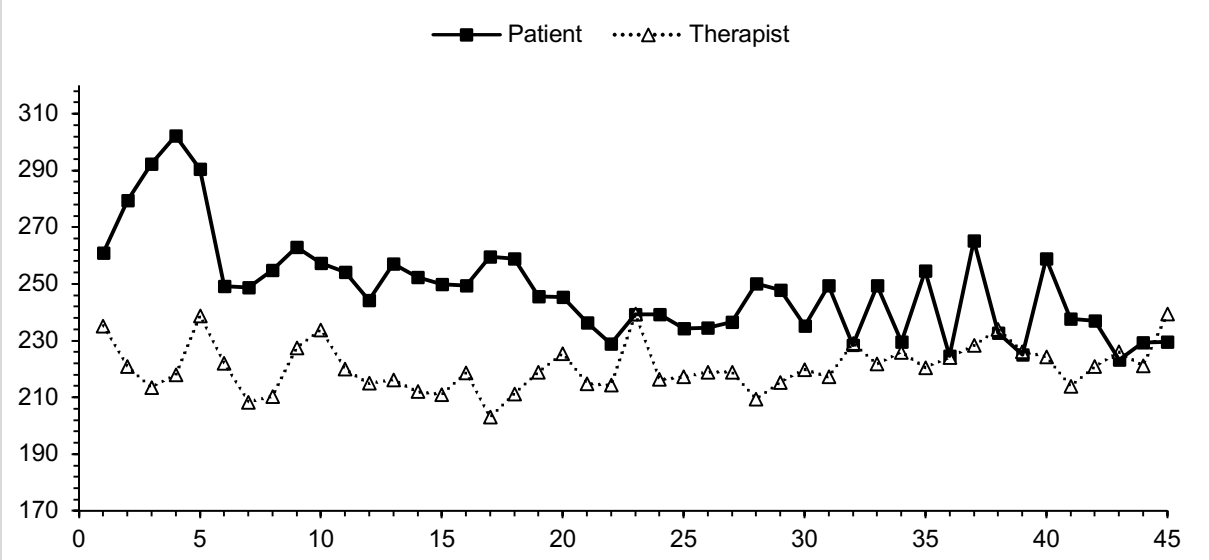


Figure 3.6. Pitch of Patient and Therapist in 6th Session



The verbal sample we'll present for qualitative analysis is selected from the time points that represent patient's maximum level of pitch. This segment also includes one of therapist's maximum levels of pitch.

Time of selection: 02:00-05:00 minutes.

P: What am I going to do? What am I going to do in the future (her voice trembles)? It doesn't seem like to me that I could succeed in marriage, either. Because I'm a person that gets uncomfortable just from sitting next to a man. I really wonder what I will do. Even if I tried to start a job, I wouldn't be able to stay at the workplace. I'm not going to school. Nothing will be the way I want it to be. I'm so scared of life. Not being able to open up to anyone exhausts me. I was all alone, and I thought about this all week. I want to cry whenever I think about these things. Because I doubt my future, I don't want to live. These things increased even more so this week; these thoughts, the feeling of not wanting to live (her voice trembles)… (Short pause). I hate; I hate myself, too. (Pause). That's it, that's my week.

T: It's like all these thoughts seem to have wrapped you up.

P: Yes. (Short pause). I don't feel happy from the things that I've done, things I used to do. Even joyous things give me pain. I'm such an incompetent, unsuccessful, powerless, desperate, cowardly person that I hate myself (sheds a tear). There are people around me who can't even access a quarter of my field of interests, somehow succeeding, managing to live. But I can't even manage that (crying). I hate myself. I used to love looking in the mirror, practicing self-care. But now, I hate that too. I want to get better but it seems that these thoughts are multiplying. I hurt so much (pause). I wish I had a sister.

In the first part of this passage, the patient talks about her fears about marriage. This looks like a disguised way of talking about the fear of sexuality. The patient is oscillating between the need to become a couple and conflicts that are triggered by intimacy and closeness. The therapist's melodic and soothing intervention, and her selection of words that "touch" the patient, is an example of her regulatory function that we observe during this session a number of times. Her intervention led the patient talk about her feelings of inadequacy and self-hatred. She also expresses her desire to have an older sister (the patient has an older brother), as if she is telling the therapist her wish to become her sister. Her words hint her need of an object of identification, which could guide her in this troublesome age of adolescence. Later in the session, the patient's depressive mood and her complaining tone abruptly disappear as she starts

to ask questions to the therapist about her personal life. Her level of pitch decreases, and her voice is filled with curiosity and playfulness at those moments. However, personal questions make the therapist uncomfortable. She responds with silence which provoke the patient's curiosity even more. During the back and forth turn rhythm that we observe in this session, it is as if the patient is trying to penetrate the veil of secrecy of the therapist. The therapist deflects these penetrative attempts and encourages the patient to fantasize about the unknown.

3.4.2. Session 8

Non-verbal theme: "Patient has a lot to say and intensity seems to be the focus"

8th session is the session where the patient has her longest total duration of vocalization (1470.81 seconds) among the six sessions. In addition, patient's speech rhythm is characterized by relatively long length of vocalizations (2.23 seconds) and short length of pauses (1.06 seconds). These two observations alone indicate that the patient have become loquacious in this session. Her speech rhythm also hints the existence of a sense of urgency. A type of urgency where the speech tempo is moderate, since the duration of vocalizations are still long, but there is little time for space. The therapist is following and matching the duration of pauses of the patient. Therapist is also following but compensating the length of vocalization of the patient. The therapist's length of vocalization is at its maximum level (2.02), and her length of pause is at its minimum (0.76) among the six sessions, mirroring the patient's changes on these parameters. This does not mean that the therapist also vocalized in large amounts. It only indicates that the therapist has matched the beat of the patient. Just like the patient, her speech included larger chunks of vocalizations and smaller bits of pauses.

This session is also the only session where we observed a cross correlation in the domain of intensity. Intensity fluctuations are led by the therapist. The patient has also its lowest autocorrelation of intensity in this session. It seems that the therapist is introducing variations in intensity (such as increasing the loudness or diminishing it), which then the patient follows. Maybe these variations that the therapist is introducing result in the patient vocalizing in larger amounts. Or, in a reverse order, maybe the therapist is trying to regulate the patient's arousal (which she senses through the "pressing" quality of the patient's speech rhythm) through introducing variations in intensity, which she seems that she succeeded in some degree.

Figure 3.7. Intensity of Patient and Therapist in 8th Session

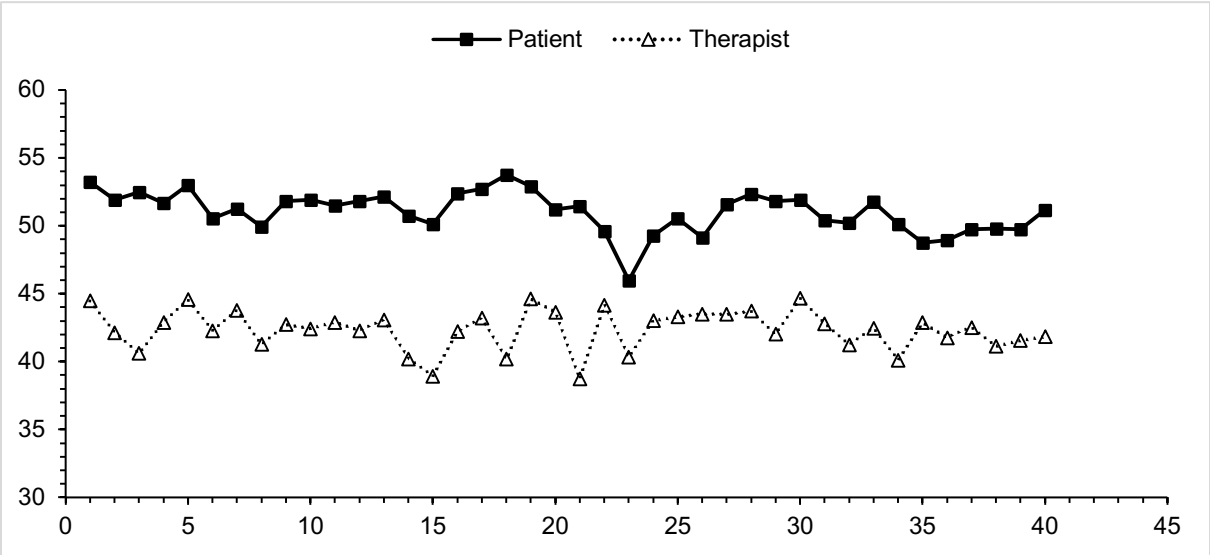
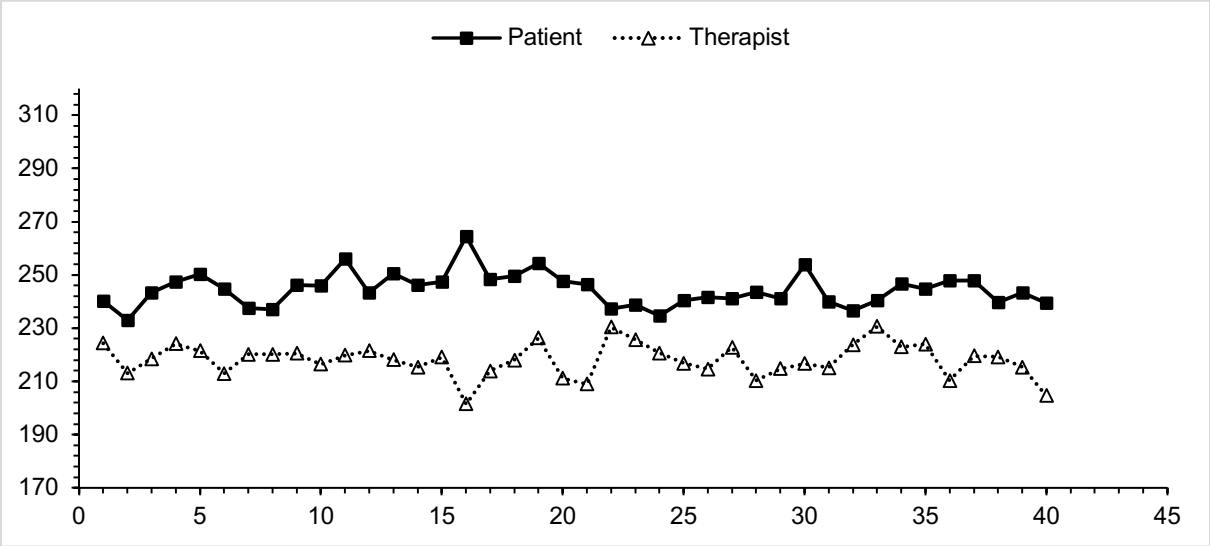


Figure 3.8. Pitch of Patient and Therapist in 8th Session



In this session, through the encouragement of the therapist, the patient starts to explore themes about her sexuality. Talking about sexuality opens up new subjects such as her fear of rape, a woman's role in society, and her desire to be a man just like her brother. It is no surprise that her speech rhythm is overflowed with vocalization: she has a lot to express. The selected one-minute segment includes patient's minimum pitch, patient's maximum pitch, patient's maximum intensity, and therapist's second maximum intensity points.

Time of selection: 15:00-16:00

P: You used to take notes, why don't you take notes now? (Laughing).

T: (Pause). I'm listening to you. (Pause). When the people outside look at you, it seems to you that they will do bad things to you. Like they are going to take you, kidnap you, rape you.

P: Yes. It's like, I have a very ridiculous life (laughing). I mean, I think about this stuff. I get scared and on top of that, I bought a book this week. I passed through Taksim to Şişli with my friend that day and bought books. And in this book, the guy at the bookstore recommended it… I started reading it; the girl gets kidnapped (T: uh-huh), goes through violence, gets raped! (T: Hmmm). And so, my fears increased. With this book, I mean, in my life, thinking that there are people who live through such things… I mean, I'm scared of being a girl, actually.

T: Hmm, of being a girl…

P: Because a guy is not exposed to these things. Even if he's all alone, no one snipes at him. No one plans to do bad things to him. If I were a man, these things wouldn't happen to me. I have fears from being a girl. Because I'm a girl. I had started life already defeated. I have to get married. I have to obey my husband.

The patient's voice in this segment is both angry and expressive. Her accentuated intensity and heightened range of pitch gives these qualities. The segment begins with her questioning the therapist in why she is not taking notes anymore. This sudden question startles the therapist and she slips into a state of discomfort, exemplified by the pauses in her speech. This discomfort possibly led the therapist impetuously restate the fears that the patient has been expressing before, in an attempt to make the patient continue to speak about herself. This excites the patient and she gets higher in intensity. Immediately the therapist senses the explosive quality of her speech as she first matches her intensity with a loud "hm hm" and then introduces a second, less intense "hmm" in order to regulate the patient.

3.4.3. Session 26

Non-verbal theme: "Therapist has things to say"

This is the session where the therapist has her longest duration of vocalization (353.34 seconds), pause (194.48) and switching pause (124.06) among the six sessions. It is also the only session where the therapist's average switching pause is longer than the patient's average switching pause. A low-to-moderate degree of autocorrelation on patient's switching pause also indicate that it is moderately predictable and stable over the course of session. Since one partner's variable of switching pause could be thought of as the response latency of the other partner, this finding indicates that in session 26, the therapist has waited less than the patient,

before initiating a turn. The fact that we see this type of pattern in only one of the six sessions suggest that something out of the ordinary has happened in this session. Beginning in this session, and continuing in the following sessions, we see that the pitch range of the therapist gets higher, becoming similar to the pitch range of the patient. Increased duration of vocalization and increased range of pitch indicate that the therapist has moved outside of her previous mode of functioning. She is becoming more active and her vocal behavior is less constricted. Also, session 26 is the only session where we did not find any cross correlation for any of the variables.

Figure 3.9. Intensity of Patient and Therapist in 26th Session

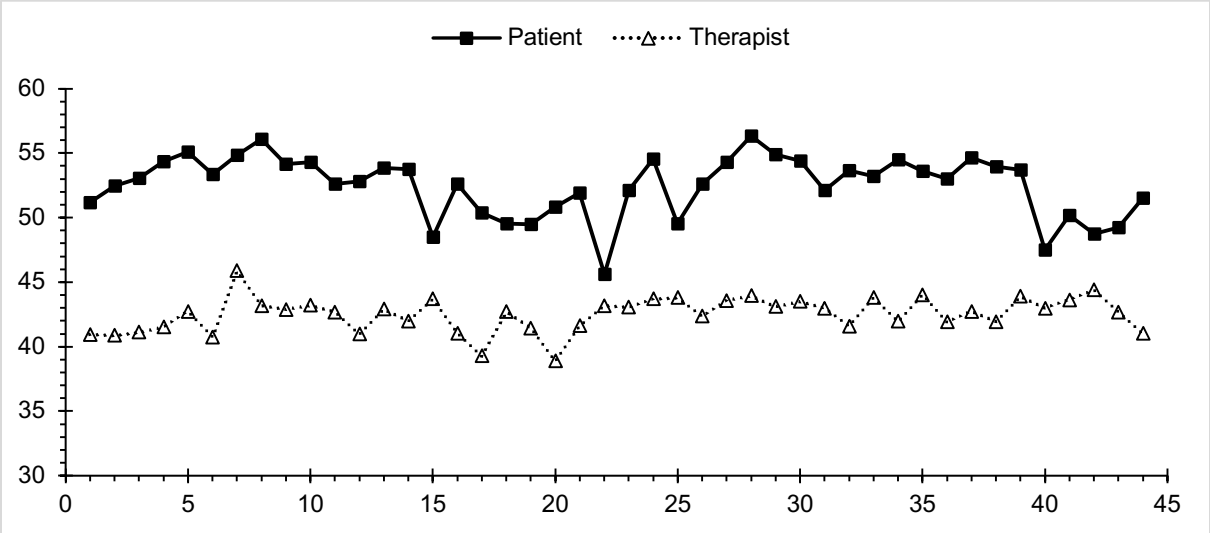
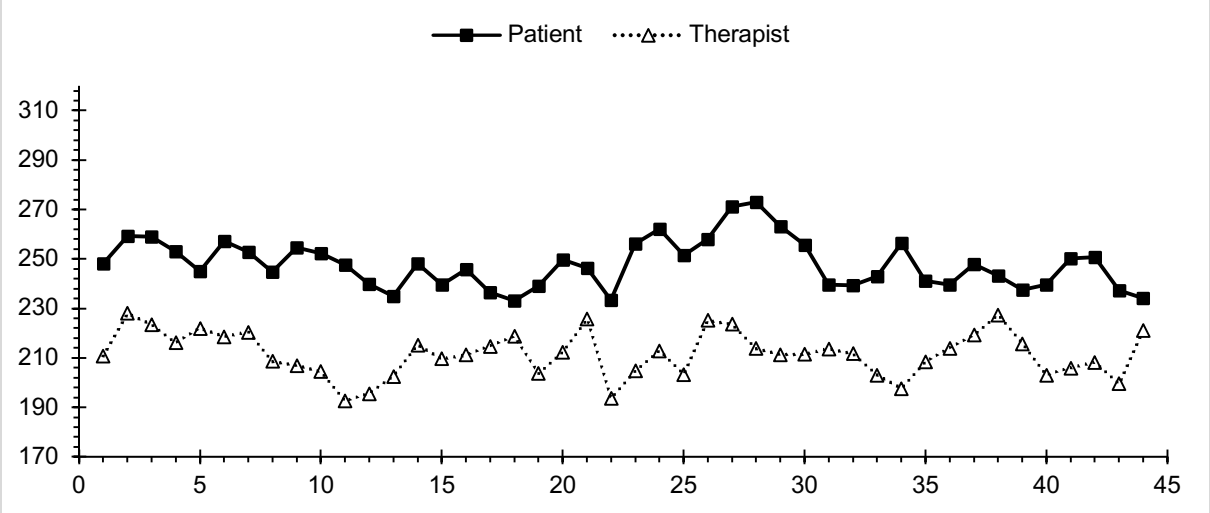


Figure 3.10. Pitch of Patient and Therapist in 26th Session



Examining the content of this session, one gets the impression of the existence of an impasse in the therapy process. Unlike the sessions from the first part of the treatment, the process feels like it came to a halt, where the same topics are discussed over and over, and there is no change. The patient is tirelessly complaining about the inadequacy she feels, and a dose of weariness is perceptible in the therapist's interventions. It is as if the therapist needs to push and shake the process, in order to open space for new experiences. The patient occasionally talks about suicidal ideas and hallucinations, which might be attempts in capturing the attention of the therapist and controlling her. The following segment includes therapist's second maximum and minimum point in pitch.

Time of selection: 20:00-22:00

P: I've started imagining things again. Like the ones I've talked about during our first sessions.

T: How?

P: I see things that shouldn't be there.

T: Like what?

P: Such as, there is no bug here. But I see it as if there is one, as if it's walking (T: Right now?) around the house. There isn't any now, it comes every once in a while.

T: When did this start? [Therapist's maximum pitch]

P: I don't know exactly. I didn't care about it that much when I first started seeing it. But then, I saw that it started happening at certain intervals. I'm scared.

T: How often does it happen?

P: Every two days. For example, it's fine at first when I go to the bathroom, but I hesitate as if I had seen something, I get scared out of nowhere. I don't know why it happens like that. I don't think these behaviors are normal.

T: Are you taking your meds as usual?

P: Uh-huh. Except for yesterday, I skipped yesterday. I fasted (It was the blessed night). I only skipped yesterday.

T: Obviously some things got pretty stirred up in your inner world. Here, the increasing number of toys also got your attention. Maybe, some things got stirred up inside your inner world, augmented. And you know, it got harder to understand them. The possible dangers seemed to have always been around. Other kinds of fears added up to your already existing fears. I wonder what are those things that build up inside you? What comes out of you when you cry and shout? Forget about who's going to say what, for now. You know, that yelling, shouting… In fact, you start seeing things that aren't there,

just so they don't call you crazy. It's because you think about how you seem. Because you think you can't explain. You keep them inside yourself.

It looks like the patient made the therapist speak in larger amounts by inducing anxiety and fear, swaying the therapist out of her regular mode of listening, where she used to follow and match the patient's vocal behavior in at least one variable and wait longer before vocalizing. Can the reversal that we observe in the average duration of switching pause be thought of as a reversal of roles? Somehow the patient has succeeded in achieving this: the therapist cannot sustain the amount of space she used to provide. In this passage, the therapist is encouraging the patient to express herself, to shout and cry. Maybe this is an invitation for both of them to break the cycle and find a way out of an impasse. The expansion that we observed in the pitch range of the therapist might be related to this.

3.4.4. Session 27

Non-verbal theme: "The emergence of a new rhythm from a quiet moment"

Session 27 has a total of 74 turns, making it the second lowest one in terms of turn number (lowest one being session 45 with 72 turns) among the six sessions. This session has the longest turn duration and longest length of pause for the patient. From this session on, patient's lengths of pauses are getting longer in relation to her length of vocalizations. The length of switching pauses are also getting longer, both for the patient and the therapist. Unlike the previous session (session 26), where the average duration of switching pause was longer for the therapist, here, we see a return to the normal pattern of switching pauses being longer on the patient's part.

It should be noted that in session 27, we observed the maximum number of autocorrelations and cross correlations in rhythmic variables. This is also the only session where we see autocorrelation of rhythmic variables in the therapist. The therapist is mostly predictable in terms of her duration of vocalization and duration of pause. It is also the first session, where we observe a cross correlation indicating the patient following and matching the rhythmic behavior of the therapist. In this session we see the highest amount of rhythmic coordination between the patient and therapist, where the therapist is following and matching the duration of pause ($r = 0.41$), length of pause ($r = 0.52$) and switching pause ($r = 0.41$) of the patient. On the other hand, the patient is following and matching the duration of pause of the therapist ($r = 0.72$). The duration of pause variable seems to be coordinated in a bidirectional way, where both members matched their rhythmic behavior with the previous behavior of the other member. The larger coefficient of cross correlation on the patient's part possibly indicates that

the patient is tracking more closely. In the following sessions, the patient continues to follow and match the therapist on this variable. Lastly, in this session the patient's strength of autocorrelation of intensity is at its highest (r = 0.46). These findings led us think that in this session, patient and therapist have found and experienced a new way of being with each other in terms of rhythm and dynamics of intensity. This experience might have provided an initial model for the long-lasting changes that we observe in the patient's speech rhythm, her length of silences becoming longer than her length of vocalizations. Indeed, in the last part of session 27, we find a period of mutual silence, with occasional vocalizations that have low intensity.

Figure 3.11. Intensity of Patient and Therapist in 27th Session

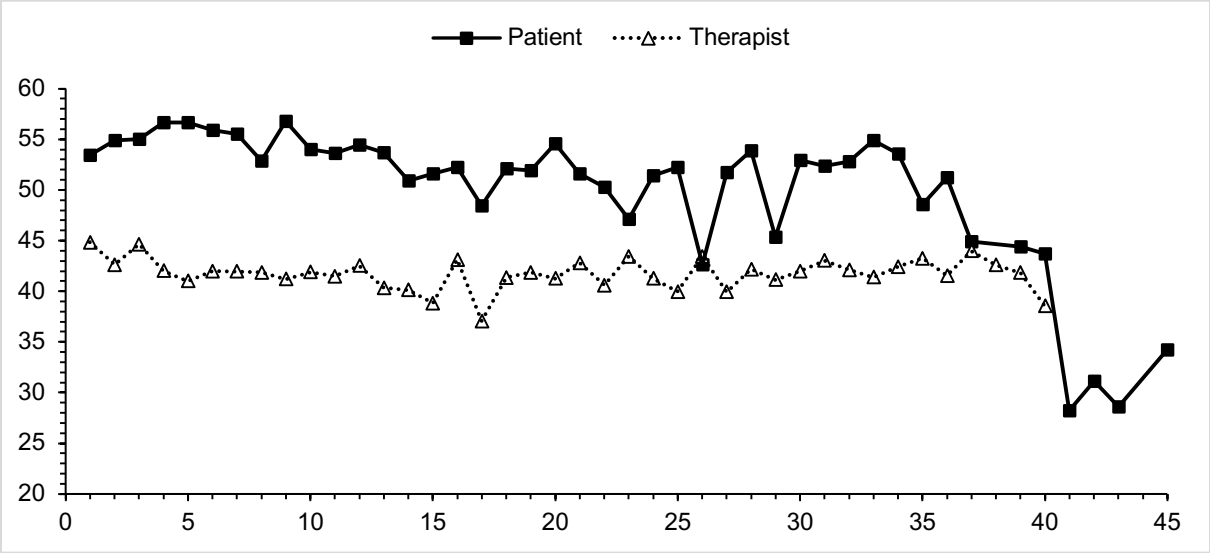
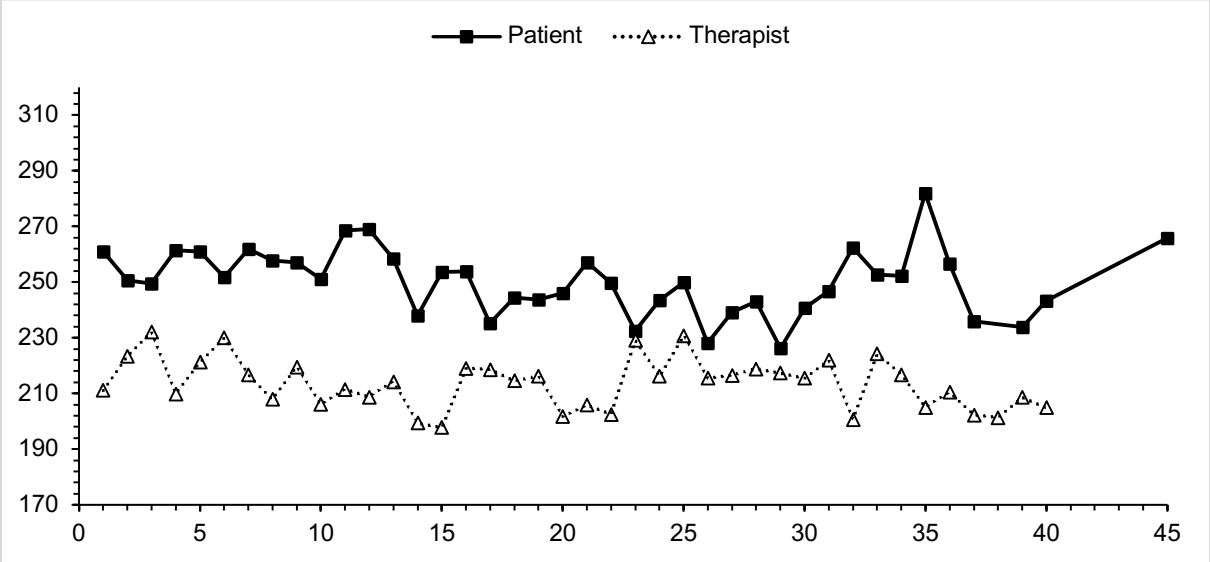


Figure 3.12. Pitch of Patient and Therapist in 27th Session



The verbal segment that we selected for this session includes the time points of minimum intensity of the patient. These are the verbal communications that are taking place at the last part of session that we mentioned above.

Time of selection: 38:30-45:30

P: How do you think I can beat my fears? When I go outside, my fears of people are very upsetting, they leave me desperate.

T: You want it to go away now. To alleviate, diminish.

P: Uh-huh. It seriously hurts. It's very hard, living scared. I don't want to go outside.

(Five minutes of silence).

P: (Looks at the watch). (Whispering) It's finished.

T: There is a little more time left.

(Silence).

[Noises coming from outside. People communicating, talking.]

P: I really wonder how people communicate with each other.

(Silence).

T: I wonder how <u>you</u> do? How you are going to? How do you want to?

Let's stop here…

Listening this segment conveys a feeling of calmness and intimacy. Moments of long-lasting mutual silence enable the voices coming from the outside to permeate into the auditory sphere of the consulting room. This creates a relational context where the patient and therapist both observe and reflect on something that is experienced together from the same point of view and is distant. Taken together with the previous session, the middle phase of treatment looks like it is characterized by frustrations, trials and errors, and new discoveries in the non-verbal domain.

3.4.5. <u>Session 45</u>

Non-verbal theme: Getting louder!

Session 45 is the session with the smallest number of turns (n = 72) and the therapist's mean turn duration is longest (13.69) among the six sessions. In this session, we also observe the longest switching pauses for both patient and therapist. Most importantly, the level of intensity and range of intensity seemed to have increased for both participants. This is the only session where the intensity autocorrelation for the patient is not present, indicating that the patient is less predictable and stable in terms of intensity compared to other sessions. From this session on, no autocorrelation is present in the rhythmic variables, indicating a similar type of

unpredictability on an individual level in the domain of rhythm. On the other hand, we observe cross correlation of two rhythmic variables, where the therapist is following and matching the duration of vocalizations of the patient, and the patient is following and matching the duration of pauses of the therapist. The unpredictability in the domain of timing, elongated pauses at turn switching and a rise in in the dynamics and overall level of intensity suggest the possible presence of a sense of threat and uncanniness.

The examination of the content of the session, reminded us the fact that the therapist has told the patient that she had to leave the country for educational purposes (in session 41). This will result in a premature termination of the treatment, and the transfer of the patient to another therapist.

Figure 3.13. Intensity of Patient and Therapist in 45th Session

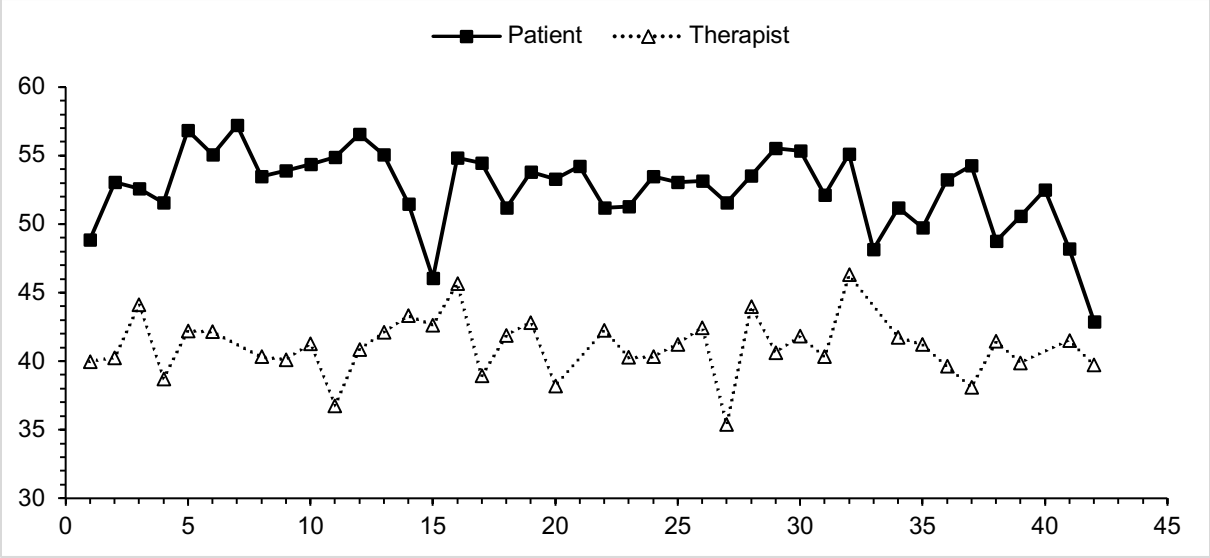
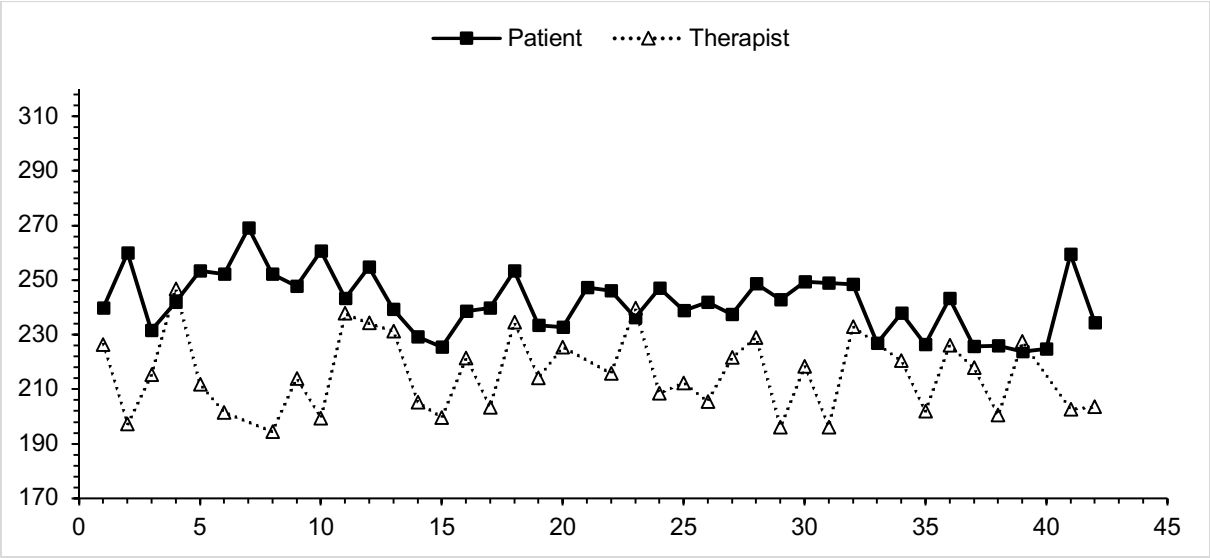


Figure 3.14. Pitch of Patient and Therapist in 45th Session.



The following sample of verbal interaction includes the therapist's maximum point of pitch, followed by patient's maximum point of intensity.

Time of selection: 04:00-05:00

P: People undermine others so much. If we had tried undermining them, we would find lots of faults. I act strangely when I enter a crowd. (Grunts). I'm also very angry. I get extremely angry. It sounds typical when I say it like this, but I really get very angry. My mother, for example, she asks something to me about something she's watching on TV, something about that; I go mad. Or my father, when he's eating, or when my mother is eating next to me, or anyone, for that matter, I go crazy by the faintest sound of slurping. This can't happen...

T: What could be making you angry? [Therapist's maximum pitch]

P: I have no idea. Normally I would have thought about it, tried to make a guess, but… I feel this taste in my mouth (demonstrates a gulp), something happens. But it makes me, makes me, very uncomfortable. I go like this: "Oh, God"! (she demonstrates: puts her hands on her head, strokes her face up – and – down). I go like this. I get crazy. After than that, I shout.

It is impossible not to notice how angry the patient is, both from the verbal communications and from the sudden rise in intensity of her voice. It seems she is angry about the upcoming separation but cannot link her free-floating anger to this fact. Her anger seems to have no object and that is how she avoids the guilt that she will possibly feel if her anger was directed to the therapist. The pauses in this session gives the idea that something is being withhold. The rhythm of speech has lost its predictability, with sudden, interrupted vocalizations that have a wide range of intensity.

3.4.6. Session 46

Non-verbal theme: Patient tracking closely the therapist.

Session 46 is the session with the highest amount of coordination between the therapist and the patient. The therapist is following and matching the length of switching pause of the patient. The patient, on the other hand, following and matching the therapist on three rhythmic parameters: duration of pause, length of vocalization and length of switching pause. The parameter of switching pause is coordinated in a bidirectional way. Furthermore, this session is the only session where we see a cross correlation of pitch and the absence of an autocorrelation of pitch for the patient. Here, we see that the patient is following and matching the therapist's pitch, alongside the other three rhythm variables.

Sessions 45 and 46 are similar in terms of the absence of autocorrelation of rhythmic variables, and the increase in average level and range of intensity. However, in this session, the patient is tracking the therapist extensively. This is the third session before the termination of treatment. The emerging non-verbal behavior might be related to this fact. Although there might be other possible explanations, the increase in matching on the patient's part led us to interpret these results from the perspective of separation and identification. Maybe the upcoming separation is driving the patient into becoming more similar to the object she is about to lose, even though she is still angry about her leaving. Or, is it possible that we see the emergence of a relational pattern where the patient is attempting to regulate the therapist? Maybe being the "following" one, instead of needing to be "followed" create a sense of mastery, which she needs in the face of approaching separation. Maybe the emergence of mimicry on the patient's part is an attempt to feel safe, confronted with an unspoken feeling of dread and aggression in the air (indicated by the existence of an unpredictable rhythm and increased intensity).

Figure 3.15. Intensity of Patient and Therapist in 46th Session.

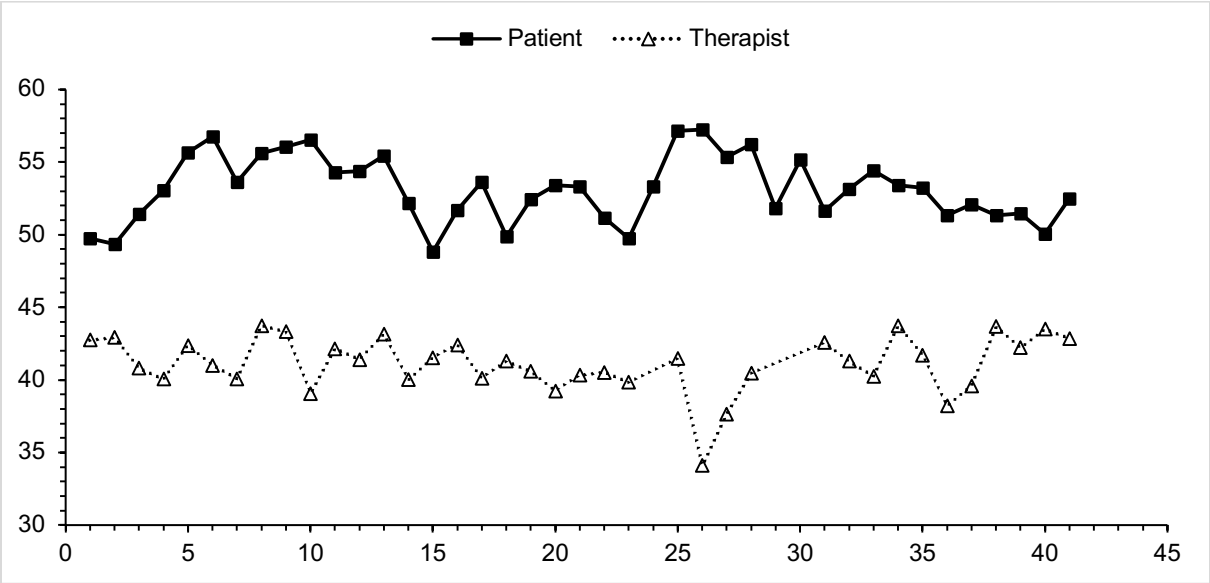
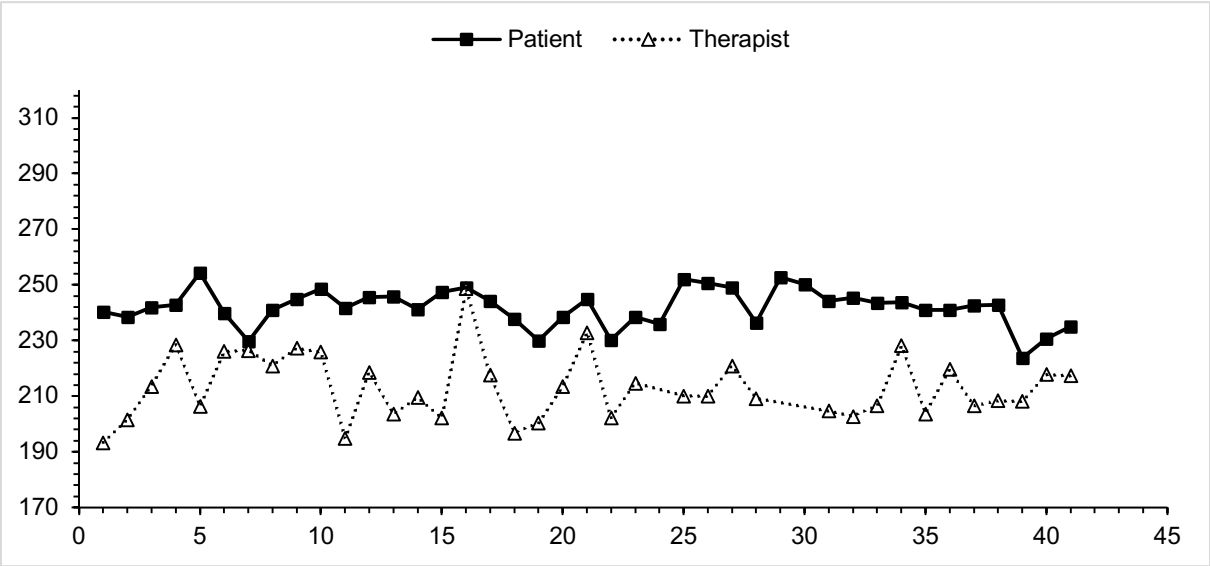


Figure 3.16. Pitch of Patient and Therapist in 46th Session



The following sample is selected according to the maximum pitch point of the therapist. Selection of time: 15:00-16:00

P: I want to shout. I want to get out and scream in the middle of this school.

T: This process with Ms. N. (her next therapist), is going to be a process where you'll be able to experience all these things you feel, all these things inside you. (Pause). It's a place where you will be able to talk about what that beginning made you live through, or what the ending will cause for you, what it already has caused, also. You can shout, you can…

P: (Suddenly, abruptly) Why Ms. T? Why!

T: …express yourself. You can do all these things.

P: I don't want to make her upset either. After all, she tries to help me, too.

T: Why would it upset her, you talking like this? [Therapist's maximum pitch]

P: I'm a little angry towards her. Anyone would have done the same thing, I don't know. I'm not going to sit still in the face of her slightest annoyance. Because I'm already… (her speech interrupts).

In this passage we see that the patient has started to forge her anger, directing it towards the new therapist that she will be transferred to. Still, the anger is not directed towards her current therapist. Taken together with the previous session, the last part of this treatment is characterized by a dissolution of predictable rhythms, increases in the mean level and range of intensity, and the patient following and matching the vocal behavior of the therapist.

3.4.7. Summary

Apart from the dominant non-verbal themes that we have presented in the preceding parts, we will now summarize some global changes that we observed over the course of six sessions. First, the therapist's turns are getting longer as the therapy approaches to an end. Also, the duration of vocalizations and the duration of pauses starts to even out both for the patient and therapist. Similarly, the patient's length of pauses gets longer, and she starts to follow and match the duration of pause of the therapist. Therapist's autocorrelation of rhythm, intensity or pitch is almost non-existent throughout the process. The therapist's vocal behavior might be determined as a result of the interaction of a large number of factors, such as her own thoughts, her considering the overall state of the therapeutic relationship, etc. On the other hand, the patient engages in constant self-regulation of pitch, intensity and rhyhtm. As the therapy approached to an end, we observed an increase in the number of cross correlations and a decrease in the number of autocorrelations.

CHAPTER 4
DISCUSSION

In this study we have attempted to depict six sessions selected from a psychoanalytic psychotherapy process of a 17-year-old patient with depressive complaints in terms of rhythm, intensity and pitch of patient's and therapist's vocal behavior. Under the category of rhythm, we have measured and reported turn duration, total duration of vocalizations and pauses in each turn and mean duration of vocalization, pause and switching pause (which we termed as lengths). The operationalization of these variables is guided by Jaffe and Feldstein's (1970) description of various temporal events of speech. For intensity and pitch, we calculated mean levels, range and standard deviation for each 30-second segment of a session. Later, we have analyzed whether each member was predictable from its earlier behavior and whether their behavior could be predicted from each other's behavior on the same parameter. For this, autocorrelation and cross correlation analyses are conducted. Therefore, we had information regarding the predictability-stability of patient and therapist's behavior on each parameter, and information on who is following the other. Combining all this information, we have identified a "non-verbal theme" for each session. Next, for each session, we have identified maximum and minimum points of pitch and intensity and went back to the original audio recordings and listened to those moments. Among those moments, we have selected the one that we thought was the best fit to the "non-verbal theme" that we have identified before. Transcripts of those

moments are presented, and the verbal content of the sessions and those moments are used in order to interpret the non-verbal data.

Through time-series analyses, for the patient, we found statistically significant autocorrelation of rhythm in sessions 6, 8, 26 and 27; intensity in sessions 6, 8, 26, 27 and 46; pitch in sessions 6, 26, 27 and 45. On the other hand, for the therapist, we found statistically significant autocorrelation of rhythm in session 27; and pitch in session 6. On certain variables related to rhythm, we found (+) lagged cross correlations (indicating the therapist is following the patient) in sessions 6, 8, 27, 45 and 46; and (-) lagged cross correlations (indicating the patient is following the patient) in sessions 27, 45 and 46. In addition, cross correlation of intensity was statistically significant only in session 8, and cross correlation of pitch was significant only in session 46, both having a (-) lag, indicating the patient is following the therapist. Qualitative analyses suggest that the early phase of this treatment was characterized by the therapist's emphatic stance that is aimed at regulating the patient's level of arousal (as indicated by high levels of pitch), instances of back and forth rhythm due to the patient asking questions to the therapist, the uncovering of topics related to femininity and sexuality and their impact on the patient's speech rhythm. The middle phase is understood as a period in which disruptive enactments, their regulation and discovery of new forms of experience are taking place. The late phase, by being colored by the knowledge of an upcoming premature termination of the psychotherapy process, is characterized by the patient's disguised anger towards the therapist, fear of separation and increased levels of mutual regulation.

4.1 COMPARISON OF RESULTS WITH PREVIOUS EMPIRICAL STUDIES

With few exceptions, most of the empirical studies that have previously depicted the temporal organization of vocal behavior are done in settings that are outside of the therapy context. In cases where the study is done within a therapy context, usually the modality of treatment was something other than psychoanalytic psychotherapy. In our review of the literature, we found out that Holtz's (2003) study was the only one that depicted the temporal patterning of psychoanalytic therapy sessions. With regard to the timing of vocal behavior, our results are similar to Holtz's findings. In her study she also found that the patient's turn duration, total duration of vocalization and total duration of pause was much longer than the therapist, indicating that the temporal structure of a psychoanalytic dialogue is different than conversations between partners in interview contexts (Cappella & Planalp, 1981; Jaffe & Feldstein, 1970; Warner et al., 1987; Welkowitz & Kuc, 1973). Lengths of vocalizations and pauses that we have measured are also similar to Holtz's findings. However, Holtz found longer

switching pause durations in two of the three cases that she used in her study. In our findings, compared to the other variables, the parameter of switching pause seems to be the one with the most variability from session to session. The comparison of descriptive results related to intensity and pitch with previous studies will be problematic since intensity is highly dependent on the technique of measurement and the mean level of pitch can highly vary from one individual to another (Rochman & Amir, 2013). Nevertheless, our observations indicate that the patient and therapist had a characteristic level of pitch that stayed stable over time whereas their range of pitch was found to be varying from session to session. This suggests that the range of pitch might be a better indicator in interpreting between session differences.

With regard to the convergence (ordinary correlation) of vocal parameters, the most consistent result that previous studies on temporal characteristics of speech have found was matching the duration of pause and switching pause, and compensating the duration of vocalization (Jaffe & Feldstein, 1970). In our study, we could not calculate correlation between sessions because of our limited sample size (n=6). However, a visual examination of Figure 3.2. suggests that the patient and therapist is matching their lengths of vocalizations, pause and switching pauses from session to session, with the exception of session 26 where the switching pause of the patient is longer than the therapist's.

With regard to the self and mutual regulation of vocal behavior, autocorrelation and cross correlation analyses revealed that first of all, the patient is mostly predictable from her behavior in terms of its level of intensity and pitch. The patient is also predictable in at least one variable of rhythm in the early and middle phase of treatment. On the other hand, the therapist vocal behavior is usually not predictable from its previous states. Instead, she seems to be regularly following and matching the patient's duration on at least one rhythmic parameter. Towards the end of the psychotherapy process, the patient also followed and matched the therapist's pitch and vocal rhythm. It is hard to do direct comparisons between our findings and previous studies (Cappella & Planalp, 1981; Feldstein et al., 1993; Jaffe et al., 2001) due to contextual and methodological variations. Again, Holtz's (2003) study is the most similar one to our study to do such comparisons. Holtz found autocorrelation and cross correlation of vocalization, pause and switching pause for both patient and therapist, consistently, in almost all of the sessions. Our findings revealed no such regular existence of autocorrelation and cross correlation in variables that are related to rhythm. We have found autocorrelation in only one or two rhythmic variables at a given session, and that variable changed from session to session. This was also the case with the cross correlation of rhythmic parameters. It should be noted that these differences are most probably due to methodological differences between the two studies. We have used a turn-based method in constructing the

time-series of rhythm variables whereas Holtz used the average duration of vocalizations, pauses and switching pauses happening within 5-second segments.

In our review of the literature, we did not find any studies that assessed autocorrelation of pitch or intensity for the patient or therapist. However, two studies (Imel et al., 2014; Reich et al., 2014) have suggested the existence of pitch synchrony, although differences in methodology and methods of statistical analyses make it hard to draw direct comparisons. We found cross correlation of pitch in only one of the six sessions, where the patient was following the pitch level of the therapist. Reich et al. (2014) also stated that in their study, therapist-leading synchrony was more prevalent. They also suggested that synchrony of pitch is negatively correlated with therapeutic alliance. Our qualitative analysis of the verbal content of the session in which we have found pitch synchrony, suggested that at that point in therapy, the patient had disguised anger towards the therapist. In this respect, our findings seem to tentatively support Reich et al.'s hypothesis that the patient might be matching the pitch of the therapist at a time she sensed that the relationship was suffering.

Previous studies have indicated the presence of an "optimum midrange" with regard to the self and mutual regulation of vocal behavior (Beebe & McCrorie, 2010; Holtz, 2003; Jaffe et al., 2001; Warner et al., 1987). Since we did not use any outcome measures or a variable that could enable us to assess the status of the therapeutic relationship at a given moment, it is not possible for us to comment on this topic. However, in our qualitative analyses, we observed an increase in autocorrelations (of therapist) and cross correlations (in which the patient is following the therapist) in sessions where we thought there are problems existing in the therapeutic relationship that needed to be addressed. In the middle phase of treatment, after following a state of non-relatedness (which we thought was the beginning of a disruptive enactment), we have observed a dramatic increase in the degree of self and mutual regulation. It might be possible that following a state of disorganization, the dyad entered in a state of vigilance, where each member was acting extra carefully, in order to find a state of balance and equilibrium. There were also increases in mutual regulation activity in the final phase of treatment. It seemed to us that, due to a fear of separation, the patient has drawn into a symbiotic merger fantasy and started mimicking the vocal behavior of the therapist.

4.2 DISCUSSION OF RESULTS FROM CLINICAL-THEORETICAL POINTS OF VIEW

Attunement, Mirroring and Containment

Our observation of the therapy process starts with a highly emotionally aroused patient. Her level of arousal is visible on the pitch graph of 6th session (see Figure 3.6.). The therapist's curve in the graph looks like as if she is mirroring the patient, but her contour moves in the opposite direction. Over the course of the session, we see that the pitch of the patient gradually decreases. Listening the session and examining the verbal contents also show that the patient's affect becomes regulated over time. The therapist's movements in pitch that are counterbalancing the patient's movements might be illustrative of a process of containment on an auditory level (Maiello, 2012).

First of all, it suggests that the therapist is attentive to the fluctuations of pitch in the patient as recommended by some early theoreticians (Reik, 1949 as cited in Templeman, 1977). Secondly, the therapist is using her pitch to regulate the patient's excitation, just like a mother does in the process of soothing the cries of the baby. Ideally, the mother/therapist provides the proper pitch contour that is in harmony with the baby's/patient's but is aimed at diminishing the excitement, providing a sound-envelope suited to the needs of the baby/patient (Anzieu, 1995; Stern, 1998b). This process could be likened to someone crafting a pottery with her hands. Like a clay that can best be shaped and given form while it is soft and malleable, this "heightened affective moment" (Beebe & Lachmann, 2002) might be the ideal moment where the therapist's regulating function can be internalized by the patient. The internalized regulatory function of the therapist will hopefully provide a model for the patient's further self-regulation of her affect. The initial sessions with an adolescent are important, in the sense that it provides the foundation of the development of a therapeutic space (Bronstein & Flanders, 1998). This space is characterized on the one hand by the therapist's attunement to the adolescent's emotions, but on the other hand, it needs to establish the boundaries of the therapeutic frame. Here, the therapist's cautious behavior in not matching the exact level of pitch seems to be effective, since otherwise the patient's affect might have escalated further. Instead, the therapist is signaling the patient the opposite direction with low pitched vocalizations; the direction where she will hopefully return following her journey (into the peaks of emotional arousal). Just like a jazz accompanist, waiting for the soloist to return to the dominant key of a track, the therapist is waiting the patient to return to her regular levels of pitch. This accompaniment,

staying at the "root" note, is a necessary element of the resulting sense of harmony and closure both in music and psychoanalytic work (Knoblauch, 2000).

We think that this process is also analogous to a parent's "ideal" behavior in a rapprochement phase, providing an anchoring point and encouraging the child for further exploration (Mahler, Pine, & Bergman, 1975). From the perspective of mentalization theory (Fonagy, Gergely, Jurist, & Target, 2002) the mirroring provided by the mother is a crucial element in the development of the notion of intersubjectivity. Fonagy (2008) argues that mirroring should have the quality of "markedness", for the infant to become able to distinguish between realness and "as-if"ness of herself and other's experiences. Here, the therapist does not match the patient in an identical manner, instead, with an implicit mirroring, opens up the space for the patient to think, creating a "reverberating" space. Birksted-Breen (2009) states that "reverberation is not just reflecting back something identical, but is reflecting something transformed by another psyche" (p. 40). Exact mirroring of pitch, rhythm, or intensity might result in a narcissistic feedback loop, where all individual differences are doomed to disappear.

The Analytic Field, Femininity and Speech Rhythms

In comparing session 6 and session 8's verbal content, it appeared to us that both sessions have similar themes related to femininity, sexuality, and relationships. But in session 6, these themes are not fully verbalized and present themselves as displacements of the original subject. Rather than talking explicitly about the fear of the sexual encounter, the patient talks about a fear of marriage. One question related to the topic of our interest is: how does this covering-uncovering of verbal material reflects itself on the non-verbal domain? Can the heightened pitch-affect that we observed in session 6 or the back-and-forth turn rhythm be related to the "yet to be emerged content" of session 8? The turn rhythm gets faster as the patient tries to learn more and more about the therapist's personal life; whether she is married, or she has a boyfriend. The patient's curiosity and the therapist's unwillingness of sharing personal details about her life created this bouncing rhythm. Before opening up, it seems like the patient wanted to know more about the therapist. It is as if something emerged in the analytic field (Ferro & Civitarese, 2015) that wasn't symbolized, and it created a relational non-verbal scene (the back and forth escalating rhythm). It was symbolized in the next session, where the patient is flooded with vocalizations.

The verbal content of the 8th session, taken together with the observed changes in the patient's speech rhythm made us think about a possible relationship between the content and the form of speech. In that session, the temporal form of patient's speech is comprised of long

vocalizations and short pauses. This form evoked the image of a "wall of sound" where her vocalizations acting as barriers, enabling only a small amount of space to exist in between. Can the temporal form of the patient's speech be linked with her conflictual feelings about femininity and her fear related to the idea of penetration? It is as if she is closing all spaces for the therapist to enter into her subjective world. The decreases we observed in the quantity of silent spaces might be due to her avoidance of a passive, receptive stance (Laufer, 1996). However, this avoidance might have been blocking possible transformations that the psyche can undergo. Winnicott (1974, as cited in Green, 1986) stated that "only out of non-existence can existence start" (p. 107). Silences have the potential to give birth to new words, new sounds and new ideas and less space means less possibility of such emergence. The anxiety that is related to a passive and stance, and the resulting dense textured "bath of words" might be one possible cause of the therapeutic impasse that we observed in the 26th session. At that point, it felt like the same topics are being talked over and over and there was no change. Is it possible that the process leading to the formation of a disruptive enactment also re-staged an already known, or fantasized scenario? At first, it looks like the patient got excited with sexual themes, and as a result, her anxiety might have motivated her in closing all gaps, hindering the possibility of psychic fertilization. The repetition of this pattern for a time might have brought the process into a boiling point, where the therapist is unconsciously drawn to penetrate, maybe forcefully, with longer and longer vocalizations (in session 26). The seductive chasing game carried the risk of turning into a rape scenario; however, we think that the therapist has managed this situation carefully and did not permit a re-traumatization. Instead, they have found – or created something else.

Disruptive Enactments, Regulation and the Emergence of New States

We will use sessions 26 and 27 in our attempt to illustrate how enactment and its subsequent regulation might give way to the emergence of new patterns of relating both on the symbolic and the pre-symbolic domains. There seems to be a contrast between the 26th session and 27th session, in terms of therapist's self and mutual regulatory activity. In session 26, we see the absence of any type of regulatory activity on the therapist's part. Contrary to her normal mode of functioning, where she is used to follow the patient in at least one variable, this session is characterized by the lack of it. We also see the patient regressing to a state where her verbal communications give glimpses of an objectless, solipsistic, one-person space, where she is talking about hallucinations. In her verbal communications the patient is reminding us, and the therapist, the existence of a previous narcissistic/omnipotent state where the object and subject

are still undifferentiated. We think that the lack of contact/regulation that we observed on the non-verbal domain is reflective of this. Relationally, this might be an attempt on the patient's part, to increase the engagement of the therapist. Maybe she sensed the therapist's detachment and by introducing topics that will surely capture her attention, tried to return to a previous state of togetherness. Her attempts might have caused the therapist vocalizing in greater amounts and introducing new verbal content to the field. The following session, session 27 is the session where we see the most rhythmic regulatory activity in the therapist, as if she is (maybe extra) carefully tracking the rhythmic patterns of the patient, also extensively regulating her own vocal behavior. The interaction in the domain of rhythmic patterning of speech gives the impression that the dyad was in a state of turbulence, swaying from one extreme to the other. When the turbulence has come to an end, we see the emergence of a new state.

From the perspective of non-linear dynamics (Galatzer-Levy, 2009; Schiepek, 2009) this type of an oscillation between the poles of chaos and order is the mechanism that brings about complex changes. One proposition of this line of thought is the argument that "chaotic dynamics enhance the capacity of a system to learn" (Kowalik, Schiepek, Kumpf, Roberts, & Elbert, 1997, p. 207). The emergence of chaoticity might be regarded as an indicator of upcoming changes. However, the system has to survive the chaos first, otherwise it might lose its form and dissolve altogether. In the last part of session 27, having survived a period of chaoticity, the patient and therapist arrive to a quiet place; a state that there is no replica of it within the other 5 sessions in our sample. Within the background of long-lasting pauses, low intensity vocalizations emerge and dissolve in a slow tempo. This also creates a permeable auditory sphere, allowing the inclusion of voices coming outside into their field[7]. The sounds coming outside of the consulting room creates a triangulated space, where a third object that is different from the patient and therapist, could be observed and talked about (Green, 1986). What is external, is framed (both spatially and temporally), encapsulated and symbolized within the boundaries of the therapeutic frame. We see that the patient is again investing on relationships, asking directly and honestly the therapist in a brittle tone of voice: *"I really wonder how people communicate with each other?"*

Freedman, Lasky and Ward (2009) stated that "enactment precedes an alteration in mental functioning" (p. 218). Their conceptualization of an "upward slope" within a cycle of transformation included phases of non-integration, transition and re-integration. In session 26, a phase of regression on the patient's part is followed by an enactment of the therapist, and this

[7] We are reminded of John Cage's famous work 4'33'' where the piece is basically composed of silence. With this composition, Cage is making a statement about the impossibility of silence, since in each performance ambient noises will fill the concert hall, creating an emergent, always changing composition.

led to a phase of reorganization in session 27. It seems that the dyad's survival of the disintegration process is rewarded by the emergence of a new, triangulated space that we observed at the end of the 27th session. Freedman et al. also identified "nodal points of change" that contained "the seeds that point the way toward the phase to come" (p. 206). The last part of the 27th session might also include the "seed" of the conflicts that will emerge in the later phases of therapy. At some later point in the treatment, the therapist informs the patient that she has to move to another country because of educational purposes. This might have created a triangular tension where the patient felt rejected by the therapist. In the last part of session 27, the introduction of an "outside" and an Other (Lacan & Fink, 2006) that the therapist's desire is directed to, are slowly setting the stage for the upcoming therapeutic work.

From the perspective of Stern et al. (1998), the quiet moment at the end of session 27 might be termed as a "moment of meeting". Stern et al. state that "a moment of meeting can create a new intersubjective environment and an altered domain of implicit relational knowing" (p. 909). In such moments they state that "the constraint of the usual implicit relational knowledge is loosened and creativity becomes possible" (p. 909). It is often argued that high degrees of self-regulation (autocorrelation) is indicative of inertia and repetition of old patterns. In session 26, (where we observed the minimum amount of autocorrelation and cross correlation) maybe the dyad needed a moment of destabilization, in order to find new avenues of being together. Immediately after the shock, we see the highest degree of regulation. The introduction of something novel might have led the dyadic system into remembering and re-establishing some of the old patterns that are already known. They have survived this turbulence, and the emerging "moment of meeting" is a byproduct of their mutual efforts. The *implicit relational knowing* might have been altered. We have previously mentioned that starting in session 27 and continuing in later sessions, the patient's "beat" of speech has been changed. Her pauses became longer in comparison to her vocalizations. This change might indicate the presence of an alteration on the procedural level. It is not possible to directly link that procedural change on the patient's non-verbal behavior to a specific "moment of meeting" at the end of session 27. However, for us, this is a good illustration of the countless processes of change that one finds in the course of psychoanalytic psychotherapy.

Termination, Mimicry and Merger Fantasies

The late phase of the treatment, sessions 45 and 46 are characterized by a decrease of patient's activity in self-regulation and an increase in the mutual regulation processes. Beebe and Lachmann (2002) suggested that the overutilization of either type of regulation mechanism

might be indicative of psychic states that are less differentiated in terms of underlying object relations. For example, a pull towards self-regulation suggests a detached, regressive, objectless state; whereas a push towards the mutual regulation might suggest a merger fantasy, where the self becomes "one" with the other. The patient's synchronizing her pitch and rhythm with the therapist in session 46 can be an indication of such fantasy. In a Kleinian sense, in order to cope with the depressive anxieties related to the loss of the object; and also, to stay away from guilt feelings resulting from her destructive wishes towards the object; she regresses to a paranoid-schizoid state of primitive object representations (Abram & Hinshelwood, 2018). In the patient's fantasy, she and the therapist became one, and her anger is directed towards the "new therapist" that she will start seeing soon. The patient's mimicry of the therapist's voice reminds an early period of fusion of the mother and the infant, where the boundaries between their bodies are not fully distinguished, with the illusion of fusion of their voices. It is as if the two voices are coming from a single source.

It is also evident that the patient idealizes the therapist. Coming from a religious, low socio-economic background, the therapist becomes a model for her. She repeatedly states that she finds her beautiful and successful. By incorporating her "voice", she becomes similar to her idealized object.

4.3 LIMITATIONS OF THE STUDY AND IMPLICATIONS FOR RESEARCH AND CLINICAL PRACTICE

The first group of limitations are related to technical issues regarding the process of recording and measurement of audio data. First of all, this study used data that was previously collected for a psychotherapy research program whose primary objective was not doing an audio analysis. The therapist who was responsible for the placement of the recording device, was not informed beforehand, and she might have placed the device at a different spot in the room in each session. This error might have an effect on our analyses, especially on the domain of intensity. The closer one gets to the microphone, the more his or her amplitude level gets, and hence, the level of intensity increases. Also, background noises can affect how people speak. In cases where there is a background noise (such as an air conditioner), individuals start speaking louder and this causes, again, the intensity level to increase. In addition, the device used in recording the audio material had a single microphone and it recorded the sounds coming from two speakers on a single file. Therefore, in order to analyze the audio for each member separately, we first needed to segment the single file into two separate tracks. This was a lengthy process and is partly responsible for the limited amount of sessions that could be studied.

Further research with technical improvements could use an audio device with two separate inputs and make sure the positioning of the microphone does not change from session to session.

A second group of limitations are related to the sample size. First of all, since this is a single-case study, our findings have limited value in terms of their generalizability. The selection of only six sessions is also problematic; it is hard to distinguish whether our observations represent a general pattern or reflect glimpses of instances that are less likely to occur. Future research could be done by increasing the number of dyads, and also the number of sessions. It would be also good to study sessions of the same therapist with two or more different patients.

In terms of this study's implications for psychotherapy research on the acoustic properties of speech, we think that the inclusion of multiple variables that are related to sound, enhances the scope dramatically. Computerized measurement of variables enables the measurement of nuances that couldn't be easily grasped by the raters, and employment of scripts makes it possible to measure large amounts of audio data with ease. However, we strongly suggest a qualitative approach in interpreting the results of computerized measurement of variables related to the "voice quality", since computer programs can fail in recognizing the contextual aspects of their outcome. A long silence, for example, can mean a resistance in some session, and in another, can suggest the emergence of an authentic way of relating. Similarly, increased levels of intensity or pitch don't always mean the same thing. In some situations, increased pitch may indicate a state of agitation, and in some situations, it occurs due to a prosodic movement related to asking a question. Finally, the auditory depiction and analysis of sessions provides one with data that could inform the sub-symbolic/implicit domain. Taken together with the verbal material, it becomes a powerful tool that can aid the qualitative analyses of psychotherapy processes research.

Clinically, attention to the pre-verbal domain, could inform about the possible state of the relationship and guide about the timing and content of upcoming interventions. In addition, intervening on the non-verbal domain itself is another possibility. Deliberate, informed usage of the musical domain might facilitate the process towards a resolution of impasses. This study also showed that changes in the rhythmic structure of a session, such as the turn tempo or a "role reversal" on the parameter of switching pause, can signal a disruption in the relational field. Clinicians might track the rhythm of the session and when they notice moments where they feel an urgency to start turns, can reflect about the possible causes of this urge. In addition, in line with some of the previous studies that suggest an optimum midrange of self and mutual regulation, this study showed that high degrees of matching rhythm, intensity or pitch might

indicate a troubled relationship. This is contrary to the common idea that therapist's empathy is transmitted through the imitation of the prosodic qualities of the patient. It is not hard to imagine a stereotypical way of speaking that is employed by some therapists, that feels overly emphatic and almost comical. Instead of mimicking the vocal quality of the patient, therapists are advised to stay in a vocal middle-ground where they keep their relation to the prosodic qualities of the patient but at the same time possess a sense of individuality. This way the therapist may introduce new melodic-rhythmic ideas that can regulate, diminish or expand the patient's ongoing vocal activity. Lastly, results of this study suggest that just like heightened affective moments, some quiet and silent moments can have a procedural impact. Clinicians are advised not to be aversive of moments of mutual silence, slow tempo and low intensity, since those moments might actually include the seed of upcoming transformations.

REFERENCES

Abram, J., & Hinshelwood, R. D. (2018). *The Clinical Paradigms of Melanie Klein and Donald Winnicott: Comparisons and Dialogues*. London: Routledge.

Anzieu, D. (1979). The Sound Image of the Self. *International Review of Psycho-Analysis,* 6: 23-36.

Anzieu, D. (1995). *The Skin-Ego*. London: Karnac Books.

Arlow, J. A. (1961). Silence and The Theory of Technique. *Journal of the American Psychoanalytic Association, 9(1*), 44–55.

Aron, L. (1990). One Person and Two Person Psychologies and the Method of Psychoanalysis. *Psychoanalytic Psychology, 7(4),* 475-485.

Bady, S. L. (1985). The Voice as a Curative Factor in Psychotherapy. *Psychoanalytic Review, 72(3)*: 479-490.

Barale, F., & Minazzi, V. (2008). Off the beaten track: Freud, sound and music. Statement of a problem and some historico-critical notes. *The International Journal of Psychoanalysis, 89(5)*, 937–957.

Beebe, B., Jaffe, J., & Lachmann, F. (2005). A dyadic systems view of communication. In J. S. Auerbach, K. N. Levy, & C. E. Schaffer (Eds.), *Relatedness, Self-Definition and Mental Representation* (pp. 23–42).

Beebe, B., Jaffe, J., Lachmann, F., & Feldstein, S. (2000). Systems Models in Development and Psychoanalysis: The Case of Vocal Rhythm Coordination and Attachment. *Infant Mental Health Journal, Vol 21(1-2)*, 99-122.

Beebe, B., & Lachmann, F. M. (2002). *Infant research and adult treatment: a dyadic systems approach.* NJ, Hillsdale: The Analytic Press.

Beebe, B., & McCrorie, E. (2010). The Optimum Midrange: Infant Research, Literature, and Romantic Attachment. *Attachment: New Directions in Psychotherapy and Relational Psychoanalysis, Vol. 4,* 39-58.

Bernholtz, B. E. (2013). The Client's and Therapist's Vocal Qualities in CBT and PE-EFT for Depression (Unpublished Doctoral Dissertation). University of Toronto, Canada.

Bick, E. (1968). The experience of the skin in early object relations. *The International Journal of Psycho-Analysis, 49,* 484.

Birksted-Breen, D. (2009). 'Reverberation Time', dreaming and the capacity to dream. *The International Journal of Psychoanalysis*, 90(1), 35–51.

Bion, W. R. (1970). *Attention and Interpretation*. London: Tavistock.

Boersma, P., & Weenink, D. (2012). Praat: Doing phonetics by computer Version 5.3.13 [Computer program]. Retrieved from http://www.praat.org/

Boston Change Process Study Group, Bruschweiler-Stern, N., Harrison, A. M., Lyons-Ruth, K., Morgan, A. C., Nahum, J. P., Tronick, E. Z. (2002). Explicating the implicit: The local level and the microprocess of change in the analytic situation. *The International Journal of Psychoanalysis, 83(5)*, 1051–1062.

Breuer, J., & Freud, S. (1895). *Studies on Hysteria*. Standard Edition, 2: 1-335. London: Hogarth Press.

Bronstein, C., & Flanders, S. (1998). The development of a therapeutic space in a first contact with adolescents. *Journal of Child Psychotherapy, 24(1)*, 5–36.

Cappella, J. N., & Planalp, S. (1981). Talk and Silence Sequences in Informal Conversations III: Interspeaker Influence. *Human Communication Research, 7(2)*, 117–132.

Cheshire, N. M. (1996). The Empire Of The Ear: Freud's Problem With Music. *The International Journal of Psycho-Analysis, 77*, 1127-1168.

Diaz de Chumaceiro, C. L. (1990). Was Freud Really Tone Deaf? A Brief Commentary. *American Journal of Psychoanalysis, 50*, 199-202.

Duncan, S., Rice, L. N., & Butler, J. M. (1968). Therapists' paralanguage in peak and poor psychotherapy hours. *Journal of Abnormal Psychology, 73(6)*, 566–570.

Erdem, N. (2005). Boş Oda. Presented at the PPPD 2005 Sempozyum, Psikanalitik Bakışlar: Aktarım-Karşıaktarım, Istanbul.

Feldstein, S., Jaffe, J., Beebe, B., Crown, C. L., Jasnow, M. D., Fox, H., & Gordon, S. (1993). Coordinated Interpersonal Timing in Adult-Infant Vocal Interactions: A Cross-Site Replication. *Infant Behavior and Development, 16*, 455-470.

Ferro, A., & Civitarese, G. (2015). *The analytic field and its transformations*. London: Karnac Books.

Fliess, R. (1940). Silence and Verbalization: A Supplement to the Theory of the "Analytic Rule." *International Journal of Psycho-Analysis, 30*, 21-30.

Fonagy, P. (2008). The Mentalization-Focused Approach to Social Development. In F. N. Busch (Ed.) *Mentalization. Theoretical Considerations, Research Findings, and Clinical Implications*, 3-57.

Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect Regulation, Mentalization and the Development of the Self.* New York: Other Press.

Frederickson, J. (1986). Psychoanalytic and auditory space. *Psychoanalysis & Contemporary Thought, 9(4)*, 641-651.

Freedman, N., Lasky, R., & Ward, R. (2009). The Upward Slope: A Study of Psychoanalytic Transformations. *The Psychoanalytic Quarterly, 78(1)*, 201–231.

Freud, S. (1900). *The Interpretation of Dreams*. Standard Edition, 4. London: Hogarth Press.

Freud, S. (1914). *The Moses of Michelangelo*. Standard Edition, 13: 211–36. London: Hogarth Press.

Freud, S. (1923). *The Ego and The Id*. Standard Edition, 19: 3-66. London: Hogarth Press.

Fuller, V. G., & Crowther, C. (1998). A dark talent: silence in analysis. *Journal of Analytical Psychology, 43(4)*, 523–543.

Galatzer-Levy, R. M. (2009). Good vibrations: Analytic process as coupled oscillations. *The International Journal of Psychoanalysis, 90(5)*, 983–1007.

Gay, P. (1988). *Freud: A Life for Our Times*. New York: Norton.

Ghent, E. (1989). Credo: The Dialectics of One-Person and Two-Person Psychologies. *Contemporary Psychoanalysis, 25(2),* 169–211.

Grassi, L. (2014). The Dimension of Sound and Rhythm in Psychic Structuring and Analytic Work. *The Italian Psychoanalytic Annual, 8*, 63-82.

Green, A. (1986). *On Private Madness*. London: Karnac Books.

Greenson, R. R. (1954). About the Sound 'MM….' *The Psychoanalytic Quarterly, 23(2)*, 234–239.

Holtz, P. J. (2003). The Self and Interactive Regulation and Coordination of Vocal Rhythms, Interpretive Accuracy, and Progress in Brief Psychodynamic Psychotherapy (Unpublished Doctoral Dissertation). Fielding Graduate University, California.

Imel, Z. E., Barco, J. S., Brown, H. J., Baucom, B. R., Baer, J. S., Kircher, J. C., & Atkins, D. C. (2014). The association of therapist empathy and synchrony in vocally encoded arousal. *Journal of Counseling Psychology, 61(1)*, 146–153.

Isakower, O. (1939). On the Exceptional Position of the Auditory Sphere. *International Journal of Psycho-Analysis, 20*, 340-348.

Jacobs, T. J. (1994). Nonverbal Communications: Some Reflections on their Role in the Psychoanalytic Process and Psychoanalytic Deucation. *Journal of the American Psychoanalytic Association, 42*, 741- 762.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C., & Jasnow, M. D. (2001). Rhythms of Dialogue in Infancy. Monographs of the Society for Research in Child Development, 66 (2 Serial No. 264)., 66(2).

Jaffe, J., & Feldstein, S. (1970). *Rhythms of Dialogue*. New York: Academic Press.

Klein, M. (1946). Notes on Some Schizoid Mechanisms. *International Journal of Psycho-Analysis, 27*, 99-110.

Knoblauch, S. H. (2000). *The Musical Edge of the Therapeutic Dialogue*. New Jersey: The Analytic Press.

Knoblauch, S. H. (2005). Body Rhythms and the Unconscious: Toward an Expanding of Clinical Attention. *Psychoanalytic Dialogues, 15(6)*, 807–827.

Kohut, H. (1950). On the Enjoyment of Listening to Music. *Psychoanalytic Quarterly, 19*, 64-87.

Kohut, H. (1957). Observations on the Psychological Functions of Music. *Journal of the American Psychoanalytic Association, 5,* 389- 407.

Kowalik, Z., Schiepek, G., Kumpf, K., Roberts, L., & Elbert, T. (1997). Psychotherapy as A Chaotic Process II. The Application of Nonlinear Analysis Methods on Quasi Time Series of the Client-Therapist Interaction: A Nonstationary Approach. *Psychotherapy Research, 7(3),* 197–218.

Lacan, J. (1966) The Mirror Stage as Formative of The Function of The I As Revealed in Psychoanalytic Experience. In B. Fink (trans.) Ecrits: The first complete edition in English, pp. 75-81. New York: W. W. Norton.

Lacan, J., & Fink, B. (2006). *Ecrits: The first complete edition in English*. New York: W.W. Norton.

Lachmann, F., & Beebe, B. (1996). Three principles of salience in the patient–analyst interaction. *Psychoanalytic Psychology, 13(1),* 1-22.

Laufer, M. E. (1996). The Role of Passivity in the Relationship to the Body during Adolescence. *The Psychoanalytic Study of the Child, 51(1)*, 348–364.

Mahler, M. S., Pine, F., & Bergman, A. (1975). *The psychological birth of the human infant*. New York: Basic Books.

Maiello, S. (1995). The sound-object: A hypothesis about prenatal auditory experience and memory. *Journal of Child Psychotherapy, 21(1)*, 23–41.

Maiello, S. (2003). The rhythmical dimension of the mother-infant relationship - transcultural considerations. *Journal of Child & Adolescent Mental Health, 15(2)*, 81–86.

Maiello, S. (2012). Prenatal experiences of containment in the light of Bion's model of container/contained. *Journal of Child Psychotherapy, 38(3)*, 250–267.

Malloch, S. N. (1999). Mothers and infants and communicative musicality. *Musicae Scientiae, 3(1_suppl)*, 29–57.

Masson, J. M. (1985). *The assault on truth: Freud's suppression of the seduction theory*. New York, NY, US: Penguin Press.

Matarazzo, J. D., Hess, H. F., & Saslow, G. (1962). Frequency and duration characteristics of speech and silence behavior durling interviews. *Journal of Clinical Psychology, 18(4)*, 416–426.

Matarazzo, J. D., Saslow, G., & Hare, A. P. (1958). Factor analysis of interview interaction behavior. *Journal of Consulting Psychology, 22(6)*, 419–429.

Matarazzo, J. D., Wiens, A. N., Matarazzo, R. G., & Saslow, G. (1968). Speech and silence behavior in clinical psychotherapy and its laboratory correlates. In J. M. Shlien (Ed.), *Research in psychotherapy*, (pp. 347-394). Washington, DC: American Psychological Association.

Milmoe, S., Rosenthal, R., Blane, H. T., Chafetz, M. E., & Wolf, I. (1967). The doctor's voice: Postdictor of successful referral of alcoholic patients. *Journal of Abnormal Psychology, 72(1)*, 78–84.

Mitchell, S. A. & Black, M. J. (1995). *Freud and Beyond: A History of Modern Psychoanalytic Thought.* New York: Basic Books.

Nacht, S. (1964). Silence as an Integrative Factor. *International Journal of Psycho-Analysis, 45,* 299-303.

Nagel, J. J. (2013). *Melodies of the Mind*. New York: Routledge.

Natale, M. (1975). Convergence of Mean Vocal Intensity in Dyadic Communication as a Function of Social Desirability. *Journal of Personality and Social Psychology, 32,* 790-804.

Niederland, W. G. (1958). Early Auditory Experiences, Beating Fantasies, and Primal, Scene. *The Psychoanalytic Study of the Child, 13(1)*, 471–504.

Ogden, T. H. (1989). *The primitive edge of experience.* Northvale, N.J: J. Aronson.

Ogden, T. H. (1999). 'The Music of What Happens' in Poetry and Psychoanalysis. *International Journal of Psycho-Analysis, 80(5)*, 979-994.

Reich, C. M., Berman, J. S., Dale, R., & Levitt, H. M. (2014). Vocal Synchrony in Psychotherapy. *Journal of Social and Clinical Psychology, 33(5)*, 481–494.

Reik, T. (1968). The Psychological Meaning of Silence. *Psychoanalytic Review, 55(2)*, 172-186

Rice, L. N., & Wagstaff, A. K. (1967). Client voice quality and expressive style as indexes of productive psychotherapy. *Journal of Consulting Psychology, 31(6)*, 557–563.

Rochman, D., & Amir, O. (2013). Examining in-session expressions of emotions with speech/vocal acoustic measures: An introductory guide. *Psychotherapy Research, 23(4)*, 381–393.

Sabbadini, A. (1991). Listening to Silence. *British Journal of Psychotherapy, 7(4)*, 406-415

Safran, J. D. (2011). Theodor Reik's Listening With the Third Ear and the Role of Self-Analysis in Contemporary Psychoanalytic Thinking. *The Psychoanalytic Review, 98(2)*, 205–216.

Safran, J. (2012). *Psychoanalysis and Psychoanalytic Therapies.* Washington DC: American Psychological Association.

Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40(1–2*), 227–256.

Schiepek, G. (2009). Complexity and Nonlinear Dynamics in Psychotherapy. *European Review, 17(2)*, 331–356.

Stern, D., Sander, L., Nahum, J., Harrison, A., Bruschweiler-Stern, N., & Tronick, E. (1998). Non-interpretative mechanisms in psychoanalytic therapy. *International Journal of Psychoanalysis, 79,* 903-921.

Stern, D. (1998b). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. London: Karnac Books.

Stern, D. N. (2004). *The present moment in psychotherapy and everyday life*. New York, NY: W.W. Norton & Co.

Strean, H. S. (1969). Non-verbal intervention in psychotherapy. *Psychotherapy: Theory, Research & Practice, 6(4)*, 235–237.

Templeman, T. L. (1977). Pitch level and pitch range as a function of client-therapist interaction in psychotherapy (Unpublished Thesis). University of Montana.

Tomicic, A., Martinez, C., & Krause, M. (2015). The sound of change: A study of the psychotherapeutic process embodied in vocal expression. Laura Rice's ideas revisited. *Psychotherapy Research, 25(2)*, 263–276.

Tustin, F. (1984). Autistic Shapes. *International Review of Psycho-Analysis, 11,* 279.

Varese, E., & Wen-chung, C. (1966). The Liberation of Sound. *Perspectives of New Music, 5(1)*, 11.

Warner, R. M., Malloy, D., Schneider, K., Knoth, R., & Wilder, B. (1987). Rhythmic organization of social interaction and observer ratings of positive affect and involvement. *Journal of Nonverbal Behavior, 11(2)*, 57–74.

Welkowitz, J., Feldstein, S., Finklestein, M., & Aylesworth, L. (1972). Changes in Vocal Intensity as a Function of Interspeaker Influence. *Perceptual and Motor Skills, 35(3)*, 715–718.

Welkowitz, J., & Kuc, M. (1973). Interrelationships among warmth, genuineness, empathy, and temporal speech patterns in interpersonal interaction. *Journal of Consulting and Clinical Psychology, 41(3)*, 472–473.

Winnicott D. W. (1967). Mirror-role of mother and family in child development. In: *Playing and reality*. London: Tavistock, pp. 111–9.

Wiseman, H., & Rice, L. N. (1989). Sequential Analyses of Therapist-Client Interaction During Change Events: A Task-Focused Approach. *Journal of Consulting and Clinical Psychology, 57(2),* 281-286.

Woo, R. (1999). Sounds of Silence: The Need for Presence in Absence. *Journal of Child Psychotherapy, 25*, 93-114.