



**HAL**  
open science

# Musical proto-narratives in early mother-infant interaction : a comparative study of French-Italian dyads

Ebru Yilmaz

► **To cite this version:**

Ebru Yilmaz. Musical proto-narratives in early mother-infant interaction : a comparative study of French-Italian dyads. Psychology. Université de Nanterre - Paris X, 2024. English. NNT : 2024PA100016 . tel-04561584

**HAL Id: tel-04561584**

**<https://theses.hal.science/tel-04561584>**

Submitted on 27 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **EBRU YILMAZ**

# **Les proto-narratifs musicaux dans l'interaction vocale mère-enfant : introduction d'un modèle communicationnel précoce dans une étude des dyades françaises et italiennes**

Thèse présentée et soutenue publiquement le 22 janvier 2024  
en vue de l'obtention du doctorat de **Psychologie**  
de l'Université Paris Nanterre  
sous la direction de Maya Gratier (Université Paris Nanterre)

### **Jury :**

Rapporteur :	M. Luca GRECO	Professeur des universités : Université de Lorraine
Rapporteuse :	Mme Iris NOMIKOU	Professeure d'enseignement supérieur : Université de Portsmouth
Membre du jury :	Mme Maya GRATIER	Professeure des universités : Université Paris Nanterre
Membre du jury :	M. Emmanuel DEVOUCHE	Maître de conférences : Université Paris Cité
Membre du jury :	Mme Rana ESSEILY	Professeure des universités : Université Paris Nanterre





## **Abstract**

Narrative has been a major area of enquiry in many disciplines in the social sciences. Narrative organization is described in everyday conversation, written text and in various art forms. Proto-narratives occurring at an early period of infancy have also been described by various authors following previous narrative studies and some additional properties peculiar to this stage of life. Many studies show that, around the second month of life, infants become active participants in social interaction. In developmental psychology, research focused especially on the timing of vocalizations that characterize infants' participation and on musical aspects, such as pitch variation and tempo, of mother-infant protoconversations. Such proto-narrative episodes have been carefully described in largely qualitative studies. The aim of this thesis is to systematically explore the narrative patterning of vocal protoconversational exchange between mothers and infants. Collections of narrative episodes were identified, coded and analyzed from 2 corpora of everyday naturalistic audio recordings of protoconversations between mothers and their 2-to-3 month-old infants from Paris and Verona. Comparisons between French and Italian samples were conducted in order to identify a minimal set of acoustic features for a rich description of protoconversational narrative. Comparisons between narrative and non narrative episodes were conducted to test the acoustic-based model of proto-narrative and to explore verbal content features. Results demonstrate that French and Italian narratives showed similar patterns in acoustic measures and similar amounts of infant participation. Results also showed different patterns of acoustic features, as well as different verbal content markers, between narrative and non narrative episodes. Furthermore, we found that infants spend more time vocalizing in narratives than in non narrative episodes. As a whole, this research provides evidence of an early communicational pattern in relaxed everyday adult-infant interaction. The implications of the narrative organisation of preverbal communication and its relation to musicality are discussed.

Keywords: narrative; mother-infant communication; vocal interaction; acoustic analysis ; infant vocalization.

## Résumé

La narrativité est un sujet qui attire l'attention dans plusieurs domaines de sciences sociales. Ce motif particulier a été exprimé sous forme de conversation, écriture et expression artistique. La proto-narrativité de la vie précoce du bébé a été décrite par plusieurs auteurs suivant les études précédentes et en prenant en compte les particularités de cette période de la vie. Autour du deuxième mois de la vie, plusieurs études montrent que le bébé devient un participant actif dans l'interaction. En psychologie du développement, les recherches ont été conduites sur la qualité de vocalisations du bébé ou bien sur la musicalité de l'interaction mère-bébé. Étant décrit qualitativement, les formes narratives n'ont pas été étudiées quantitativement. Dans cette recherche présente, nous voulons étudier systématiquement comment cette narrative forme se produit dans les protoconversations vocales entre la mère et le bébé. Les enregistrements vocaux des seize dyades de Paris et de Vérone ont été analysés pour cette recherche. Cette analyse nous permet d'identifier une collection minimale des caractéristiques acoustiques pour décrire des narratifs protoconversationnels. Ensuite, les mesures acoustiques et le contenu verbal du discours de la mère ont été étudiés en prenant en compte plusieurs aspects indiquant la forme narrative des épisodes narratifs et non-narratifs. Les résultats suggèrent que la comparaison des épisodes narratifs des italiens et français sur les mesures acoustiques et participation du bébé présente une forme similaire dans les deux cultures. Les patterns des épisodes narratifs de l'échantillon français diffèrent de ceux du groupe de non-narratif sur les mesures acoustiques et le contenu verbal. De plus, la contribution du bébé dans l'interaction est plus significative dans les épisodes narratifs. Dans l'ensemble, cette recherche suggère les prémices précoces d'un pattern interactionnel; une forme de pré-organisation qui prend forme autour de la musicalité partagée.

Mots-clés: narrativité; communication mère-bébé; interaction vocale; analyse acoustique ; vocalisation du bébé.

## **Acknowledgements**

I would like to thank,

Maya Gratier for being with me during this story of my life.

Manuela Lavelli and Heidi Keller for generously hosting me in their respective universities as visiting student, especially Manuela Lavelli for sharing with me the Italian data collection for my research.

Emmanuel Devouche for his insights on the results of my study and for being so helpful.

I would like to express my gratitude to mothers and infants from Paris and Verona. I would like to thank Guilietta Von Salis and Anne Bobin-Bègue for the collection of the French data.

All my friends from Team C14 for keeping company during ups and downs. Especially, Julien Magnier who was with me at the beginning of this journey and shared the thrills of studying this difficult but exciting concept.

Finally, my family; my mother Beyhan, my dear grand-parents Yılmaz and Özden, my uncle Kayhan and my aunt Emel, and finally, all my friends around the world.

# Table of Contents

<b>Abstract</b> .....	<b>3</b>
<b>Résumé</b> .....	<b>4</b>
<b>Acknowledgements</b> .....	<b>5</b>
<b>Introduction</b> .....	<b>9</b>
<b>Chapter 1. Narrative structures: From text to speech</b> .....	<b>13</b>
1.1. Written narratives .....	14
1.2. Oral narratives .....	16
1.3. Ethnopoetics .....	20
1.4. A cognitive look at narratives .....	22
1.4.1. Narrative experience .....	22
1.4.2. Narrative organization .....	23
1.5. Proto-narratives .....	25
1.6. Narratives as structure or process .....	27
1.6.1. Agency .....	28
1.6.2. Temporality .....	29
1.6.3. Ways of expression .....	31
1.6.4. The self .....	32
1.7. Towards characterizing proto-narratives in early vocal interaction .....	34
<b>Chapter 2. Contexts of early mother-infant vocal interaction</b> .....	<b>36</b>
2.1. Early encounters through sound .....	37
2.2. Post-birth vocal interaction .....	39
2.2.1. Infant directed speech .....	39
2.2.1.1. Melodic properties .....	41
2.2.1.2. Signature tunes .....	42
2.2.1.3. Verbal content .....	42
2.2.2. Infant-directed singing .....	44
2.2.3. Infant's vocal repertoire .....	45
2.2.4. Turn-taking .....	49
2.3. Early socialization practices .....	51
2.3.1. Language socialization .....	51

2.3.2. Parenting styles .....	54
2.4. Towards conceptualizing narratives in early mother-infant vocal interaction .....	57
<b>Chapter 3. Communicative Musicality: We are all born musicians.....</b>	<b>59</b>
3.1. Communicative Musicality in mother-infant interaction .....	59
3.2. Primary intersubjectivity and interactional synchrony .....	64
3.3. Neurocognitive roots of musicality .....	67
3.4. Musical narratives: Saying something musically .....	69
3.5. Pursuit of characterizing musical narratives.....	73
<b>Chapter 4. Meaning: At the crossroads of words and melodies .....</b>	<b>78</b>
<b>Chapter 5. Methods .....</b>	<b>87</b>
5.1. Rationale and aims of the study .....	87
5.2. Research questions .....	88
5.3. Sample selection from the EVOC corpus .....	90
5.4. Sample selection from the VERONA corpus .....	90
5.5. Criteria for identifying vocal interaction sequences in the EVOC and VERONA samples.....	91
5.6. Indicators for creating collections of narrative episodes from the EVOC & VERONA sample... ..	92
5.7. Criteria for identifying non-narrative episodes (EVOC sample only) .....	93
5.8. Collections of narrative and non-narrative episodes .....	93
5.9. Acoustic analysis of the collections .....	94
5.9.1. Coding of vocal units .....	94
5.9.2. Acoustic measures .....	94
5.9.3. Infant vocalization .....	95
5.10. Narrative analysis of verbal content of maternal utterances .....	95
5.11. Acoustic analysis based on fundamental frequency and intensity graphics .....	96
5.12. Hypotheses .....	97
5.13. Statistics.....	98
<b>Chapter 6. Results .....</b>	<b>99</b>
6.1. Study 1: Narrative episodes in French and Italian samples .....	99
6.1.1. Duration of French and Italian narrative episodes.....	99
6.1.2. Narrative indicator chart.....	101
6.1.3. Principal component analysis .....	103

6.1.4. Comparison of Vocal Units in French and Italian Narrative Episodes .....	106
6.1.4.1. French intra-group comparisons of narrative episodes based on vocal units .....	106
6.1.4.2. Italian intra-group comparisons of narrative episodes based on vocal units .....	108
6.1.4.3. Comparison of vocal units between French and Italian narrative episodes .....	110
6.1.4.4. Comparison of acoustic measures between French and Italian narrative episodes ...	114
6.2. Study 2: Comparison of vocal units in French narrative and control episodes .....	116
6.2.1. French intra-group comparisons of control episodes based on vocal units .....	117
6.2.2. Comparison of vocal units in narrative and control episodes within French data.....	117
6.2.3. Comparison of acoustic measures between French narrative and control episodes ....	121
6.2.4. Narrative analysis of verbal content in narrative and control episodes .....	126
6.3. Fundamental frequency and intensity graphics of narrative and control episodes.....	130
6.4. Case examples .....	132
<b>Chapter 7. Discussion.....</b>	<b>142</b>
7.1. Discussion of the results of the present study .....	142
7.2. Discussion of narrative organization through the lens of early vocal interaction .....	150
7.3. Limits of the study .....	160
7.4. Perspectives for future research .....	161
<b>Chapter 8. Conclusion .....</b>	<b>163</b>
<b>Synthesis in French .....</b>	<b>165</b>
<b>References.....</b>	<b>180</b>

## Introduction

Over the first year of life, infants achieve necessary steps in socialization through rapidly changing communicational skills, which are thought to develop in close relation to the day-to-day interactions with social partners, particularly with parents.

Parents speak to their infants in specific ways, with pitch variation, heightened rhythm, and repeated melodies and speech content. Indeed, when engaging with them, parents often tell stories to their infants. These stories can have various forms across cultures, such as songs, lullabies, lived or fictional stories, and nursery rhymes.

Since Aristotle's *Poetics*, narratives in the text have been described by numerous authors, with relative distinctions in their structural design. Furthermore, contemporary researchers have conducted more formal analyses of narrative structures in speech and everyday conversations in narrative studies. Similarly, several specialists have portrayed music as a narrative introducing the idea of a nonverbal narrative. Given this recent evolution in how academics in various fields describe narrative structure, there is reason to empirically explore the range of characteristics that identify spontaneous narrative discourse in early social interaction.

This dissertation investigates the idea that *narrativity, or the process and ability to create narratives*, is a central basis of human cognition and socialization. We have chosen to define this interactional process using the term “proto-narrative” to describe the units of expressions created by the mother and the infant, emphasizing the earliest form of a pattern that generalizes to a broad variety of communicative narrative formats. The proto-narrative form in early interaction may contribute to cognitive developmental processes in the first years of life. The concept of proto-narrative takes its roots from observing the early processes of exchange and attuning between mothers and infants in interactional formats, namely in the work of Colwyn Trevarthen and Daniel Stern.

We are interested in the insights from a description of early social interaction between infants and parents as “Communicative Musicality”. In vocal interactions, we



can find verbal and nonverbal components to analyze. We wish to investigate whether and how narrative structure appears in these informal communication frameworks. We are intrigued by the idea that narrative acts as a common thread connecting everyday verbal exchange and the multiple ways in which art is expressed worldwide.

The two studies presented in this dissertation probe the proto-narrative organization of everyday social interaction between mothers and young infants. Together, they aim to describe how mothers' talk is organized according to the frame that defines proto-narrative sequences and how infants partake in these organized sequences, in turn shaping their own communicative development. We focus on changes in the acoustic features, such as pitch and duration, of both mothers' and infants' vocalizations during face-to-face interaction. The first study explores various acoustic features of proto-narrative vocal interaction between French and Italian mothers and their infants. Comparison of dyads from 2 similar European cultures attempts to define a minimal set of common sound-based indicators that describe proto-narratives and could be applied to other languages and cultures. The second study aims to further test the validity of acoustic as well as content-based indicators by comparing previously identified proto-narrative sequences with a random selection of sequences matched for duration, IDS quality and infant participation.

Narrative is considered a powerful format for cultural transmission. If some of the earliest interactions between infants and adults have a narrative form, long before infants can make up and share stories through language, it would follow that infants be acquire cultural habits from a very young age and before joint attention's ratchet effect (Tomasello, Kruger & Ratner, 1993). Could cultural experience permeate social interaction through a narrative organisation of expressive behavior?

During my Ph.D. research, I had an opportunity to travel to Italy as a visiting researcher in Manuela Lavelli's laboratory at the University of Verona, where I had access to work on her database of Italian mother-infant dyads. We chose to work on these two European cultures with strong commonalities in values, social organization, and representations of infancy and languages that share lexical and syntactical features and rhythmic organization. This allowed us to focus on universal practices for cultural

transmission, practices we think may be rooted in communicative musicality. Likenesses, along with contrasts, may provide insight into cultural practices and parenting styles as well as reflect uniqueness in terms of individual, social and cultural peculiarities. Observing the interactional styles of infants and mothers from two different cultural groups may provide information about cultural transmission, social roles, shared practices, and cultural continuity.

We chose to study interactions of 2-to-3-month-old infants because developmental studies have shown that this period is crucial in terms of awareness and vocal expressiveness of the baby. Parents start to perceive their infants' vocalizations as communicative and meaningful at these ages.

Language acquisition and socialization are essential markers for typical development and possible developmental deviations during the first years of life. Many studies of productive language development begin from the babbling phase, for its discernible syllables are easier to transcribe and study from a methodological point of view. However, we propose that the early communicative skills of infants prior to the babbling stage, such as vocalizations, are also important to take into account and give us significant clues on the future nature of developmental aspects.

My own experience in clinical psychology and linguistics, both as a practitioner and researcher brought me to work with children diagnosed with autism spectrum disorder. I intended to create and show children with autism that there is another way to communicate than words, and that, I thought, was the role of music. Seven years ago, with my colleagues, we created a musical environment for these children using musical instruments as well as our own voices as instruments. In a semi-improvisational framework, we organized our therapeutic sessions to have a well-defined beginning and end, allowing children to express themselves through music. This experience made me more curious about the roots of communication. This led me to pursue a Ph.D. in early communication to better understand its forms and functions from a developmental perspective. This doctoral dissertation is based on the earliest stages (2-to-3 month-old) when infants start to express themselves toward the outside

world. The main purpose of this study is to discern a pattern that may be the basis of all later communication, reflecting both its individual and cultural foundations.

## Chapter 1: Narrative Structures: From Text to Speech

« Sous ses formes presque infinies, le récit est présent dans tous les temps, dans tous les lieux, dans toutes les sociétés ; le récit commence avec l'histoire même de l'humanité ; il n'y a pas, il n'y a jamais eu nulle part aucun peuple sans récit; toutes les classes, tous les groupes humains ont leurs récits, et bien souvent ces récits sont goûtés en commun, par des hommes de culture différente, voire opposée : le récit se moque de la bonne et de la mauvaise littérature : international, transhistorique, transculturel, le récit est là, comme la vie. » *Roland Barthes, Introduction à l'analyse structurale des récits, 1969, p:1.*

The Latin word “*narrare*” refers to telling, or narrating. Until recent times, narrative studies were exclusively literary. Most research on various narrative styles began in the twentieth century in domains such as cinema, music, or media. Various methodologies have been applied to analyze different forms of narrative in psychology, sociology, and linguistics. More recently, multidisciplinary approaches have been developed to study narrative works and processes. Literary narratives rooted in language provide more solid ground for detailed analysis than nonverbal narrative forms such as dance.

Plato emphasized the distinction between *diegesis* and *mimesis*, the performance of telling by the narrator and the performance of the characters by enacting a story. Some of the first constructed reflections on the narrative structure can be attributed to Aristotle. In his *Poetics*, Aristotle first examines the relationship between reality (life) and the work of art. Art is the imitation of reality, described as *mimesis*. He illustrates three modalities of mimesis: medium, such as music and language; agent or character, such as the object of imitation; and narrative or drama, such as a sequence of imitated events. He was interested in the genre of tragedy and emphasized a typology of characters and series of events leading to a cathartic effect.

For Aristotle, “The Mythos” or the plot, is the principal part of the tragedy and comes before “the character.” The plot is the “imitation of the action,” which is “the arrangement of incidents” representing the action.

### **1.1. Written narratives**

Jumping to the 20th century, we see that many experts found inspiration in Aristotle’s analysis of narrative and opened new perspectives on narrative research. Russian folklorist Vladimir Propp was interested in the narratives of folk tales and studied a corpus of Russian tales. In “Morphology of the Folktale” (1928), he formalized the typology of narrative structures in which he distinguished actors from actions. He focused on narrative features, the relationship within and between, and the whole narrative rather than on a narrative schema. He portrayed “functions” as minimal units representing a significant action in the intrigue and a position in the timeline. According to his theory, several functions combine chronologically to form a plot. Propp identified thirty-one functions in the plot structure of fairytales. Furthermore, Propp’s intensive analysis highlighted the structured schemas and their functionalities within the narrative process.

French Structuralists contributed new research on narrative analysis in the sixties with the emergence of semiotics. Barthes (1969) accentuated the universality of storytelling, saying that if there is life, there is a story. Overall, the narrative structure is described as beginning with an initial situation presenting an introduction to the subject while the hero’s status remains stable. An initiator (trigger) then modifies the initial situation, introduced by a temporal connector or incidents, in other words, twists and turns provoked by the trigger, and followed by an outcome, leading towards the end and final situation (e.g., Todorov, 1968). On the other hand, different levels of narratives have also been described through minimalist indicators.

Influenced by Propp, Lévi-Strauss studied the structures of myths (1958). In a similar vein to Propp’s functions, he describes the “mythèmes” as fundamental propositions of the myth, describing the relationship between subject and function. According to him, these “mythèmes” combine into ensembles to create meaning.

In “Les catégories du récit littéraire” (1966), Todorov presents two levels of categories: *story*, which is fictional, and *discourse*, which comes up in a specific context depending on interlocutors, time-space, and linguistic components (Todorov, 1978). For Todorov (1968), narratives draw the following path: they start with a state of equilibrium that will be disrupted (disequilibrium), and then this disruption will be recognized and repaired. This series of events will be the outcome of a new state of equilibrium, similar but not identical to the initial one. Conjointly, he describes two types of episodes within the narrative: the state (equilibrium and disequilibrium) and the transition between states.

Genette (1972) presents three different descriptions of “narrative” in contemporary usage: the first one points to the *narrative statement* as a means to tell an event in oral or written discourse; the second one is attributed to describing *narrative content* which is related to the dynamic between the discourse and the events; the third one indicates the acts generating the discourse.

In “Actantial Analysis,” Greimas (1969) distinguishes “récit,” i.e., a discourse describing actions, from “narrativité,” which is an organization of meaning that all discourse retains. For Greimas, narrative grammar appears as a modality to decrypt the organization of meaning expressed in utterances. Two major elements emerge: *actants* (agents), state or transformative utterance units, and *transitions*, shiftings from one state to another. This succession gives the narrative its temporal dimension, a relationship between what happened before and after, which causes “the inversion of the situation.” Todorov (1971) additionally mentions “the transformation of the content,” which corresponds to Greimas’ assertion, to which he adds narratives of zero-transformation, ones where no switch situation has been observed.

Along with the temporal link between these series of events, a causal bond between events is also indispensable. For Ricoeur (1983), the consistency of the whole narrative relies on how it takes shape around the plot. With chronological succession between events, narratives must also be organized to become meaningful wholes.

Mostly, what we observe in text or theater two forms of storytelling, which can be described as time travel in discourse. The first includes a narrator who tells the story

from a third-person perspective and describes one or multiple protagonists who act within the story. In the other form, the narrator and the story's protagonist are the same, and narrating occurs through living the story. Thus the events are perceived from a second-person perspective. Barthes (1953/1972) brought two essential functions of writing into focus: the use of the *passé simple* (literary past tense in French) and the third person. The French narrative past tense no longer exists in spoken language but is only used in writing and is a narrative's linguistic and artistic indicator. The narrative past is used to indicate the past and represents a temporal spectrum, linking sequences of events and maintaining their causality. For Barthes, the third person is also a narrative convention, acquainting the reader with the plot and assisting the reader in creating a narrative world without forgetting its fictional nature. Various, Genette's description of a narrative is based on a minimal parameter. For him, a single utterance describing that *something happens* is sufficient to describe a narrative, and this is through the "expansion of a verb" (1972, p.75). The echo of that straightforward happening creates the narrative effect. Usually, these actions have a beginning and an ending; they carry out temporal proceedings between these two points based on the action's affordances (Genette, 1972).

## 1.2. Oral Narratives

This section will explore discursive and conversational forms of oral narratives by reviewing sociolinguistic and linguistic anthropology studies. Oral narratives appear to be a broad term covering different approaches to understanding its dynamic formation. This section covers the organization of the story, the impact of linguistic characteristics, and the social and cultural aspects of oral storytelling.

In 1964, Dell Hymes introduced the term "ethnography of speaking" to study how people talk in particular contexts, a term he later changed to "ethnography of communication" to widen the linguistic frame of his study. He proposed six basic units to conduct the ethnographic analysis of communication (Hymes, 1972): *speech community*, people sharing the same communicative rules; *speech situation*, the place in which the scene takes place; *speech event*, an event with a beginning and an ending in a speech situation; *communicative act*, any act carrying out the meaning;

*communicative style*, particular personal speech styles; and *ways of speaking*, styles or patterns belonging to a culture.

From a sociological perspective, Goffman (1956) framed everyday-life face-to-face interaction through metaphorical concepts. The dramaturgical framework he presents reveals how the person presents himself or his activity, the impression he makes on others, his *performance*, and the impact of the environment on his performance. Charles Goodwin (1984) analyzed the organization of telling a story through participants' organization and their place in the story, ways of interpreting the components of the story, and alternative ways for the action to be enacted through body and talk. Additionally, he studied the active participation of story recipients by asking questions that invoke the affective and constructive process during storytelling activities (Goodwin, 1986).

To understand their complex structure and for a deeper analysis, oral narratives needed a different approach, a more flexible one compared to the structural narrative analysis of a text. American sociolinguist William Labov was interested in the oral narrative structures of personal experience, such as how people narrate their lived past events (Labov, 1972; Labov and Waletzky, 1967). Labov & Waletzky defined oral narratives as devices that helped summarize past experiences by matching successive events through verbal propositions. Based on collected data from citizens of New York in the presence of an interviewer, they labelled certain levels to describe the build-up of their discourse. The organization is broken down into the following components: it starts with an abstract (a synopsis of what comes next) and is followed by indicators (who, when, what, where), by a complicating action (development), an evaluation (how the narrator evaluates the action), a result (resolution) and a coda (outcome of the events). According to Labov (1978), narrative propositions can be distinguished from free propositions because their temporal design presents a chronological ordering of events.

There is, thus, a natural selection when it comes to telling stories. We pay attention to narratives because they have been significant events for us, different from everyday routines that are predictable. They are therefore events worth telling (Labov,



2001; Bruner, 1991). Labov (2006) insists that only the development part is essential to identifying a narrative because it points out why this event is worth telling. Van Dijk & Kintsch (1983) also mention that an exciting and vital event should be expressed because the “complication” caused by an unpredictable event merits telling. This aspect appears to be the main criterion distinguishing narrative from other forms of discourse.

Inspired by classic narrative analysis, Labov adapted the form to everyday conversations from their collected data to provide a framework for oral discourse within narrative studies. He gathered some critical elements based on temporal organization, inner structure, evolution, and evaluation within the narratives and inserted the listener's perspective as well as his own assessment. Labov noted the importance of the review of the story and suggested “reportability” as one of the main criteria defining a narrative. The narrative, for him, is a fundamental form of communication. One of his most important contributions to narrative analysis is to have determined the “line of dramatic tension” (1972) as the affective flow within the story. This authenticity makes a story compelling.

This framework is appropriate for most oral narrative situations, such as therapeutic interviews, oral memoirs, and, most importantly, oral narratives of everyday conversations. In the case of therapeutic interviews, selective memory is also involved in narrating the most significant events and also calls for associated events to resolve the problem coming from the initial event.

Yet another essential feature of storytelling is the engagement of the attention of the listener. This medium depends on how the *raconteur* tells the story and how the auditor reacts to the story's details. These details rely on different styles, verbal content, or in some cases, variations of acoustic patterns of voice. These qualities reveal the degrees of the emotional upheavals and transformations within the story for all parties involved. However, in Labov & Waletzky's work (1967), the focus is on the verbal content, in other words, what the teller brings up in their personal story.

Linguistic anthropologist Elinor Ochs has a similar view to Hymes (1981) in terms of a performance-based approach, emphasizing the importance of the audience's

feedback during the storytelling process. For Ochs (1997; 2002), narratives manifest in our knowledge and belief systems, values and ideologies, actions and emotions. It is a common pattern we use to think about and interact with our past, present, and future. Furthermore, stories are seen as cultural tools supporting the child's socialization. In terms of expressing narratives, verbalization is not ubiquitous for Ochs. She also mentions different forms of narrative expressions, such as musical, kinesthetic, and pictorial forms and embodied gestures and facial expressions.

Essentially, Ochs suggests that the emergence of narratives has a collective dimension in social life. She analyzed a large corpus of ordinary talk and family stories. Her studies (Ochs, 2011; Ochs & Shohet, 2006) propose that narratives are not only an individual communication style but also a social activity that implies different participants and roles. The audience is always involved in the story together with the narrator. Interpretation and orchestration of the narrator and reinterpretation of the audience allow the community to represent events and emotions together to create a common sense and approach to resolve problems and instantiate personal and collective identities.

In "Living Narrative" (2002), Ochs & Capps describe an approach to narrative analysis through five dimensions. First of all, they underline the importance of sharing the story. The story belongs not only to the one who narrates but also to the audience; feedback and reactions of the audience and recall of their memories related to the story. Like Labov's "reportability," telling significant events that make an impression on the interlocutors is also essential to capture their attention and react. In the case of an uninterested event, this dimension depends on how the narrator presents his story rhetorically. Another point is the relation between the story and the current ambiance: is the story related to the ongoing or previous conversation topic, or is there a rapid subject change? They also mention the internal organization of the narrative, expecting a logical or linear organization of the sequences. Finally, the last dimension is the story's impact on the conversation partner herself.

Together with Labov, Ochs & Capps (2002) also suggest those narrative events are more "problematic" and effective in people's lives than other events of everyday

life. The particular unforeseen event gives a plot effect, captures the listener's attention, and creates "dramatic tension." Ochs & Capps evaluate narrative practice along two axes: the first includes the problematic event and its resolution, called the narrative structure; the second is based on the aftermath of the told event.

While Ochs & Capps emphasize the importance of sharing the story with others, Hanninen (2004) specifies three different modes of individual narrative: told, inner, and lived. The inner narrative represents the interpretation of past, present, and future events of an individual's life. It becomes told narratives once the individual decides to share them and filters them as they are told. Lived narratives, accompanied by inner narratives, are situations in which we live everyday lives and impact our actions.

In "Story and Discourse," Chatman (1978) describes narratives as combining what is told and how it is told. Story resides in two elements: *events* (what happened) in both active (actions) or passive (happenings) manifestations and *existents*, characters living the event, and the place where it happens. Discourse is the way the story has been exposed and revealed to us. Genette, in "Narrative Discourse" (1972), specified three points to help distinguish between story and discourse based on temporal properties: the order chosen to tell the events, the repetition of an event, and its duration.

### **1.3. Ethnopoetics**

Following his studies on the ethnography of communication, Dell Hymes later introduced an ethnographic approach to analyze spoken narratives, called "ethnopoetics," using corresponding poetic structures in speech. This approach defines oral narratives as culturally-influenced aesthetics of speech. Hymes first mentions ethnopoetics, the performance-based narrative analysis, in his study of the Chinook community from north-western America (Hymes, 1981). In the language of the Chinook, Hymes identified the repetition of specific themes along with the repetition of syntactic structures and suggested that this community has a peculiar system of linguistic structure. Hymes formulates that the narratives of Chinook people are comprised of scenes referring to the participants of the narrative and the relationship

between them, which are represented by acts referring to a fragment of an episode in the narrative.

Later, in his work entitled “Ethnopoetics and Sociolinguistics” (1996), Dell Hymes revisited Labov’s (1972) narrative theory and reanalyzed it through an ethnopoetic perspective. He suggested that oral narratives are formulated and transcribed through verbal units and formulations, lines and groups of lines, verses, and stanzas at different levels. Meaning-making is organized and mediated by diverse levels of verbalization enclosed in acts, which Dell Hymes calls “poetic patterning” (2000: 320). These levels and their relationship contain a dynamic similar to metrical equivalence in poetry. He accentuated the fact that stress, intonation contours, and high accent need special attention while analyzing verses in terms of verbal art similar to poetic forms.

Hymes mentions “parallelism” (Jakobson, 1960; Hymes, 1981), indicating the relationship among verses within the spoken narrative in terms of recurrent lexical and syntactic sequencing through repetition and variation, such as repetition of words (rhyming in some cases) or topic. Hymes defines this as an internal organization of each line of the narrative and suggests that it is metrical, like in poetry (Hymes, 2000). The poetic metricality comes with this parallelism and displays both linguistic and cultural traditions. Jakobson (1960) emphasizes a poetic internal structure, proclaiming verbal expression as a work of art. However, Hymes also observes an external “measured” regularity, an organization of relations between the lines of a spoken narrative, preceding the internal organization within the lines of a spoken narrative as poetry (Hymes, 2000, pp. 295-296).

This organization gives a pattern of relationship between lines of the spoken narrative. The relationship between parts of a spoken narrative, either metrical or measured, carries out one’s personal and traditional storytelling style. However, Hymes describes more than one “narrative grounding” to tell the same story depending on the expressive choice of the storyteller (2000: 313). This kind of deviation in patterning is also reflected in the internal organization of verses, suggesting a metrical shift. Hymes reports that along with patterning, there are elaborations, dramatically changing the narrative flow, that he calls amplification

(Hymes, 1985: 412). The amplification indicates a shift within the recurrent verse pattern; it is “a level within the narrative level” (2000: 301).

Ethnopoetics views oral narratives as performance rather than discourse. Hymes offers a view of ways of speaking, including ethnography of communication, as an alternative to narrative analysis focusing only on the internal structures. He considers the relationship between distinct levels that he describes internally and externally. Through these relations, he suggests a new approach for analyzing speech organization in terms of metricality, voiced in different degrees, reflective of cultural and individual patterns. Hymes puts forward the significance of adopting a broad view of linguistics to thoroughly seize the transmission of cultural and linguistic knowledge, or as written in terms of “the narrative voice” (1996).

#### **1.4. A Cognitive Look at Narratives**

This section aims to transition between the linguistic and psychological realms, reviewing cognitive studies investigating narratives. It mainly centers on the narrative experience and organization embedded in two parts: the narrative experience is examined in a tripartite way, including body, mind, and the environment; in other terms, how lived experience shapes narratives and the narrative organization of the human mind is covered on the cognitive basis, in light of temporal organization, representation, and memory.

##### **1.4.1. Narrative Experience**

A growing body of research on human consciousness from cognitive psychology and neuroscience brought narratives into focus. Indeed, inspiration originates from phenomenology, the physical embodiment of our lived experiences. As phenomenology focuses on the subjectivity of experience, it claims that the human body is inseparable from the world (Merleau-Ponty, 1962). This view has inspired many researchers and propagated new research under the name of embodied cognition. Cognition, here, is regarded as interactive and physically manifested.

Aligning with this perspective, Lakoff & Johnson (1980) put forward conceptualizing abstract concepts through metaphors. Metaphor, at this point, refers

to a cognitive representation rather than a linguistic characteristic. These metaphors lean on some previously attained basic mental concepts that stemmed directly from physical experience. These basic concepts can be thought of in parallel with affordances (Gibson, 1979), which are cognitive actions that the environmental aspects provide, the perceived action, and even future possibilities that the initial one bears. Metaphor is the transfer of a literally experienced concept to an initially unknown representative concept. Lakoff & Johnson suggests that the physical experience of the body with the environment prepares first-draft meanings for more complex concepts later on. The interaction of these basic concepts or representations progressively sets the forthcoming linguistic meaning (1999).

From a more neuroscientific point of view, neuropsychologist Antonio Damasio (1999) studied the neuronal basis of cognition, emotion, and the self. His research focuses on human consciousness and how the mind responds to the body's reactions to the world through experience. He describes three dynamic levels of self by defining the self itself as a dynamic process (1999). The transition between different levels, first automated, then representative, and finally complex, accompanies how we transit from primary representations to a more elaborate personal life story. This process starts as nonverbal narratives before it becomes a verbal flow, by acts of representation propagating from both internal states and external objects. Implicit storytelling is a natural process created by the organism through numerous levels (Damasio, 2010).

#### **1.4.2. Narrative Organization**

The previous section covered how experience shapes our representations and how these representations are dynamically improving and expanding. This section alternatively focuses on narratives' internal organization and their impact on different cognitive-linguistic levels.

The temporal organization appears as one of the major characteristics of narratives in many forms and research. But priorly, it is suggested that this temporality is pre-organized before becoming expressed via different ways of storytelling. The oral narrative involves some form of temporal contingency between successive sentences.

Sentences follow and complement each other in chronological order, thus contributing to the process of meaning-making. At least two temporally ordered sentences should be consecutive to identify narrative sequences (Labov & Waletzky, 1967; Labov, 1978).

In 2006, Labov revisited his past work on narrative structures to underline a new emphasis on narrative pre-construction. Together with other authors interested in how narratives organize or reorganize our memory, Labov accentuates anew that the narratives are pre-constructed before they are told. He indicates that telling narratives requires a cognitive evaluation to select the reportable event and establish the temporal order within which the narrative will be told to the audience as clearly as possible. Recalling memory for telling a past event is affected by time. Therefore, some transformations are recognized over time between the lived situation and when the event is being told (Labov, 1997).

For Ricoeur (1983), the chronological order of the events is the leading property defining narrative. There is a strong connection between narrative function and the human experience of time. He distinguishes physical time from narrative time, which is principally “human time.” A person, for Ricoeur, is the character who builds up his own internal narrative and shares it with others later (1990).

Brès (1994) suggests that oral narration is designed to provide meaning to the told event; in other words, he contends that we do not tell an event to communicate information but to share the meaning of the event with the conversation partner. On this principle, he revisits Ricoeur and Labov’s assessments of temporal organization and affirms that time in oral stories does not correspond to real-time order; the teller follows an order primarily by choosing a significant moment of the event first. While living the situation implies the progression of the events, telling the event does not require the exact chronology. Basically, the real-time temporal occurrence of actual events is transformable while reciting; what matters is to make sense of it, and changing the order of events does not affect the relevance.

In psychology, the temporality window is also seen as an essential shaper of cognitive organization and memory. For Bruner (1990), narratives are fundamental formats for social interaction while forming the basis of our cognition. This basic

cognitive capacity is already constructed socially and culturally even before being linguistic. He supports the idea that our experiences gathered in our memory are organized in narrative formats. These narrative formats expose the self in terms of reflections, speech, and description of the real world in a diachronic way. The process of reality construction or making sense of the world has a narrative manner of organizing past knowledge and experience along with memory and social and cultural cues. Bruner views the temporal frame in broader terms describing narrative diachronicity as an “account of events over time” (Bruner, 1991: 6). He emphasizes that time in narratives should be considered parallel with its meaning. Bruner assigns a lengthwise approach to time, which he describes as “human time,” like Ricoeur (1983), rather than sequential time.

Bartlett’s early work (1932) on memory resonates with Bruner’s proposition of a narrative basis for the organization of selves. For Bartlett, stories are organized in order, and events are positioned within the sequences of “introduction-development-conclusion.” Memory thus appears as a process of narrative reconstruction. Bartlett developed “schema theory” to describe abstract mental representations organizing events in a structured way.

Along the same lines, Roger Schank portrayed “scripts” (Schank, 1975; Schank & Abelson, 1977) as mental structures developed through experience. Scripts lead to basic units and then unfold into expanded and enchaind dynamics of our memory, such as plans and themes. In “The Literary Mind” (1996), Mark Turner also mentions that narrative function is, first of all, a cognitive capacity, first created as “small stories” before being shared through speech.

### **1.5. Proto-narratives**

Many researchers in psychology working on young infants’ worldviews and others working on emotions and behaviors in social interaction argue that interpreting perception or even experiencing events is narrative by nature. The following part focuses on researchers studying mother-infant interactional settings to understand how they conceptualize proto-narratives as a shared experience at the very beginning of life, before access to language. These authors adopt a similar perspective upon



describing narratives as lived experiences; their body of research is built on clinical observations and empirical findings of dyadic interaction.

The idea is undoubtedly developed alongside the work of Jerome Bruner, who focuses on cognitive learning, language development, and narratives considering socio-cultural aspects of interaction. Bruner (1990) centers his research on how the mind creates meaning by emphasizing observations in everyday life. He also revolved around the cultural impact on the narrative shaping of ourselves and the social world within which we live, together with language acquisition. Bruner (1991) concentrated on human beings relating to and reciting our lives by nature, principally how we organize and interpret our everyday experiences and the diachronicity of language.

Recognizing prelinguistic forms, Bruner's highlight on language acquisition of narrative thinking does not limit the subject matter to verbal narratives. According to Daniel Stern (2000), whose center of attention is mainly on infant development, lived stories are organized narratively in the infant's mind, represented in the form of an affective movement, and are non-verbalized. He illustrates the "present moment" as an experience-based moment incorporating a beginning, an ending, and a plot. Temporal contours represent shifting time along and within the lived experience (2004). He introduced the concept of "proto-narrative envelopes" (1985/2000, 1992) as a story form without words, framed by a temporal structure providing meaning. Although this structure represents the earliest form of a narrative, it nevertheless contains an agent, a context, an intention, and an action. Proto-narrative envelopes originate in early interpersonal contexts by viewing themselves and the other during constant interaction and widen through affective tuning or "attunement" of the mother-infant duo. These early proto-narrative envelopes spread out grounds for the infant's narrative self and identity and are coupled with the culture they belong to (Stern, 1992, 1999, 2000).

Colwyn Trevarthen's extensive research and collaborations are also based on the psychological development of infants and define experiencing time as central to progress (Trevarthen, 2008). Trevarthen (2000) explains that even from birth, infants are capable of taking an active part in a dialogue by taking turns expressing themselves

and are sensitively receptive to the mother's expressions. Together with Stephen Malloch, he developed the theory of "Communicative Musicality" (Malloch, 1999; Trevarthen & Malloch, 2002; Malloch & Trevarthen, 2009), which illustrates many aspects of the nonverbal dimensions of the interaction between a mother and her baby, and brings about the analogy between communicative functions of music and language. Their body of research consists of the existence of a language without words, its musical nature, and the psychological origins of musicality during mother-infant interaction across multimodal segments of analysis such as melody, rhythm, and movement. The narrative nature of early vocal and nonverbal interaction constitutes a significant part of their thinking. These early narratives rapidly turn into ritual-like settings of their genuine interaction where the mother and the infant enjoy creating meaningful moments (see Chapter 3 for further details on this theory). The narratives are not only limited to vocal expressions but are multi-modal with body movements and facial expressions, fully arranged in synchrony during this early meaning-making process (Delafield-Butt & Trevarthen, 2015).

### **1.6. Narratives as structure or process**

This section aims to cover narrative studies of different expressions, such as written or oral, through reviewing studies of diverse authors and approaches to examining narrative units and structures. Written narratives allow researchers to conduct formal analysis based on steady compositions of the text. However, an oral narrative requires different methodological formulations due to its interactional form. This interaction is not only observed between the teller and the audience, but it also exists between internal components of the story. The interactional aspects of narrative have bound multiple disciplines together and encouraged researchers to conduct further analysis. Therefore, to bring this chapter to a close, we will proceed through a question: do we consider narratives as structure or process? We will answer this question by reviewing the common characteristics of narratives and positioning interaction as a primary principle affecting both sides.

### **1.6.1. Agency**

A first notice could be made upon the narrator and/or the agents of the story in light of the distinction between diegesis and mimesis in Ancient Greek Theater. In the first case, the narrator is the third person telling the story without taking part and presenting an external content frame to the audience. On the other hand, the agents or performers are the ones who live and participate in the act of telling through multimodal ways of expression. Both parties are performers; mimetic performance stages an interpersonal interaction; however, the diegetic performance offers an interaction between the story and the narrator. In the scope of written texts, the latter is seen between the author and the text, as well as the reader and the text. This contrast raises more questions in interactional storytelling settings, especially in everyday life, where multiple agents contribute to the process. As we have seen in the examples of sociolinguistic studies (Labov, 1972; Labov & Waletzky, 1967), the narrator is the teller and the one who experienced the event told. On the other hand, the audience is not only composed of the spectators but also of the active participants of the story's formulation through the medium of their comments or questions (Hymes, 1981; Ochs, 2011; Ochs & Capps, 2002).

Further interrogations can also be made upon telling and living or experiencing a story. Here comes the question of temporality regarding the time lapse between real experiences, which later becomes a verbal story. Last but not least, narrative as a shared experience in interaction is also worth considering (Stern, 2004; Trevarthen, 2002).

The difficulty of conceptualizing different agencies in narratives obviously arises from the study of oral narratives by means of assigning more than one position to the participants. This difficulty directly impacts the dynamics among communicational partners, constantly changing the internal structure of the narrative; however, the external structure, such as the topic, remains the same.

Those issues influence the methodology of narrative analysis as well. Labov & Waletzky (1967) significantly contributed by including everyday storytelling in narrative research. However, while collecting their data, they assigned an interviewer as a

conversational partner who asked questions about the significant events they lived to the subjects they studied. This method was later criticized by Schegloff (1997), saying that the study of ordinary storytelling has to preserve its natural co-constructed and collaborative process without interrupting it. Many developmental psychologists followed Schegloff's footsteps in their analysis, maintaining the natural flow of early mother-infant interaction.

Bruner (1990) attributed great importance to narratives, shaping our reality as a predominant discourse form existing even before the language. Bruner's research focuses on the structure and functions of narratives and observational studies of how children recite their everyday lives (Bruner & Lucariello, 2006). This article presents a more profound view of how a young child narrates her recent past life in monologues. This monologue has a metacognitive function of constructing her own world through the medium of language. Yet again, the agent here shifts and casts the child as a storyteller of her own story without the presence of a visible audience.

Roles and participation of different agents transform the storytelling process and generate diverse narrative structures, particularly in the case of oral narratives in interactional settings. In that case, narratives appear to be an interactional process rather than a steady structural entity. However, the internal content shares joint components such as temporality.

### **1.6.2. Temporality**

Another common component of narratives impelled researchers, alternatively named temporality or temporal organization. Many researchers described the organization based on temporality as units that build up the internal structure of narratives. These units, however, were labelled in various ways depending on the focus.

The classic textual analysis defines narratives as the succession of three main parts: introduction, development, and conclusion. The further detailed analysis put forward more micro-units and questions the order of presentation.

Labov & Waletzky's analysis (1967) suggests that the oral story opens up with a foreshadowing of the plot, then gives a context indicating time, place, and agents, and

finally presents more details about the story, including tension (Labov, 1972). The question of the plot was also argued among researchers. Hymes (1996), reviewing Labov's work, proposed that a number of oral narratives do not fit Labov's model. On the same wavelength, Fludernik (1996) puts forward the existence of narratives without a plot in the study of oral narratives; however, she exclusively underlines the importance of a human experiencer to indicate "emotional involvement" and its cognitive mirror (1996, pp. 9-10).

Many others also studied the urge for sequentiality. Bruner emphasized the importance of sequentiality in how we construct our world. He pointed out narrative diachronicity precisising the occurrence over time and its "irreducibly durative" nature (Bruner, 1991: 6), which is not enough to point out a narrative without its canonical structure. There he introduced "breaches" within the canonical order, which makes the story worth telling to distinguish narrative forms from others (Labov, 1978; Bruner, 1991).

Fludernik (1996) proposes an approach to view narratives as a part of social experiencing and suggests "experientiality" as a core process to emphasize cognitive parameters' leading role similar to text interpretation. This one indicates a deviation from traditional narrative modelling and instead focuses, first of all, on real-life experiencing. She accentuates that the storytelling process "operates as a cognitively grounded frame." (1996: 11). This view can be thought of in parallel with embodied cognitive studies mentioned above, in addition to the narrative expression of social interaction.

Narrative as an embodied human experience also finds its place in mother-infant interaction. The emphasis on "present time" and the nature of engagement, along with shaping shared emotions, is the focal point of attention to investigate this prelinguistic early narrative formation (Stern, 2000, 2004). Temporality here is seen as similar to a musical phrase for its internal flow and narrative-like sequences, as well as the outer limits of beginning and ending (Stern, 1992). In other words, the temporal frame is experienced as "infant-time," a temporal envelope shorter than any other form of oral narrative, with an internal dynamic conveying the meaning before the speech.

Temporal inquiry does not only points out the temporal organization of successive events but also sequences within the narrative structure. In the intersection of experience and order, the question is to understand or describe narratives, which brings up different grounds, including different modalities. The temporal organization also raises several questions around defining narrative coherence; for some, it is strictly attached to temporal sequentiality, but others question the necessity of its existence.

### **1.6.3. Ways of Expression**

The review of research on narrative studies revealed that narratives are expressed in many ways: texts, discursive, conversational forms, and earlier forms of the prelinguistic period.

Analysis of verbal content covers the majority of research in narrative studies dating back to written narratives. In the examples of Labov (1972) and Bruner & Lucariello (2006), oral expression of the verbal story allows us to investigate internal structures that shape a narrative. However, verbal content is not isolated from complementary expressions such as intonation and gestures. Verbal and nonverbal expressions usually interact without disregarding purely nonverbal expressions such as dance. Particularly in early interaction, the musicality of speech, communicational gestures, and facial expressions are essential aspects of expressing ourselves (Delafield-Butt & Trevarthen, 2015; Malloch, 1999; Trevarthen, 2002). Researchers like Bruner (1990) and Ochs (1997) also suggested that the way we express ourselves carries cultural and social cues. Cultural-specific motifs also point out peculiar aspects of expression from the culture and society we belong to (Hymes, 1981, 2000).

These different modalities of expression highly influence whether the narrative is a structure or a process. While oral storytelling of a past event owns a sketch of the narrative in mind, spontaneous and interactive forms reflect a co-creation during the conversational flow. From this perspective, various expressive forms offer us different levels of narratives.

#### 1.6.4. The Self

Narratives also have a symbolic function in how we shape our identity through social interactions. In the early stages of life, emotion derived from synchrony between communicational partners and its representational functions provides the infant a ground to discover the self through the other. Narratives are the way of scaffolding identity through early communication. The early sketches of self are conceived through meaning-making processes mirrored by the conversation partner. These are essential steps contributing to the infant's vocal and cognitive development (Bruner, 1985, 1990).

De Jaegher & Di Paolo (2007) suggest that sense-making in early interaction takes shape through the mother's intention to provide meaning to her infant during their interaction. Authors call this process orientation, the adjustments of meaning between communication partners through coordination. The mother assists the infant through the way during the exchanges to make sense of the shared experiences.

Sharing stories with others is the communicative function of narratives. Before being shared and told, those stories are pre-constructed in our minds (Labov, 2006). Delafield-Butt (2018) points out that narratives between mother and infant are co-constructed "in lived time" and become written or composed later. First, stories are lived or enacted, and then, they are told or written. Enacted stories have symbolic functions beyond words of discourse (Langer, 1942). These representations are seen as the first steps in modelling how we regulate our emotions (Delafield-Butt & Trevarthen, 2015; Porges, 2011), leading to recognizing and conceiving parts of ourselves.

Studies on musicality show how important indicators the musical features are to genuinely seize both individual and interactive expressions and their unique traits. Margaret S. Barrett (2011) conducted a three-year longitudinal study to investigate how music-making shapes "an emergent identity" musically and culturally. Inventing songs and music afford narratives and allow children to perform and express themselves through music, reflecting their unique selves over musical expressions. Signature tunes (Bergesson & Trehub, 2007; Gratier, 2001) could also be seen as part

of this process of conveying self-expression. Musical signatures are directly linked to the uniqueness of individual vocal expression and are part of our identity.

Special attention is also dedicated to rhythm and synchrony to comprehend better the role of mutuality in interaction and its consequences on the infant's state. Brazelton, Koslowski, and Main (1974) say that the acts of adaptation between mother-infant dyads are related to the repetitive-rhythmic organization of this interaction. Another example of this mutual adaptation is Condon & Sander's (1974) research published in the same year, observing the synchronized leg movements of infants' and adults' speech. In similar studies from a musical standpoint, Corbeil, Trehub & Peretz (2016) suggest that infant-directed singing's repetitive and rhythmic patterns postpone distress reactions of the baby. Babies prefer happier sounds over neutral sounds (Corbeil, Trehub & Peretz, 2013). Trehub, Ghazban & Corbeil (2015) reveal that infants' contented or neutral states last longer when listening to a play song than infant-directed speech. These studies display evidence of how the distinct musicality and regular rhythm could effectively maintain and regulate the affect.

According to Nelson (1993), narratives are crucial to the development of the self, an autobiographic self, and a meaning-making process of a child's inner world and environment. Different levels of self were also described by Damasio (1999). The protoself, preceding consciousness, is the self-operated governance of the organism. The coreself is crafted through stories by experiencing different levels of perception. These levels consist of representations of internal states such as memory or external objects. Even at this level, the story in the form of a nonverbal narrative has a protagonist, and the process is constantly dynamic, waving from one representation to the other. The autobiographical self, an extended version of the coreself, is consisted of past memories allowing future projections without discarding the present. Swinging between the past and its influences on anticipation of the future and conserving the present leads the self to think in terms of narratives. The extended version manages to create a life story owing to editing our personal stories of life (Damasio, 1999).

Stern (2000) describes "senses of self" through his clinical observations of children and their caregivers. He describes four stages of development of self:



preverbal self, intersubjective self, verbal self, and narrative self. Stern says the transition between stages is not conclusive; there is a progression through stages; nevertheless, that does not mean that all features of previous stages disappear entirely. The preverbal stage is the skeleton of what comes next, the intersubjective self, which is the co-shaping of the consciousness. This leads to the verbal self, which is a basic form of the narrative self that the baby attends at the age of three. This model shows the stages an infant passes through to reach the narrative self, yet the self is present before language.

Bartlett (1932) says remembering is justifying an attitude and has a constructive function. Recollecting patterns from the past serves as a memory reconstruction for our interlocutors and us. Markus & Nurius (1986) describe past and present selves along with possible selves that we can project into the future. As we connect repetition with the past, we can surmise that possible selves are related to the variation, inserting a new pattern derived from a familiar repetitive pattern. Elaborating repetitive patterns along with the now self helps introduce new patterns, and “new” represents the future. Sociologist Maurice Halbwachs’ work (1925) mainly focuses on how we manufacture our collective memory. For him, narratives are the social texture of our memories; social groups contribute to the common life and share stories. Besides, they abstract a common interpretation of this storytelling experience. From this point of view, Halbwachs recognize narratives as a product and the embodiment of social structure. Ricoeur (1983) additionally emphasizes the developmental nature of narratives to constantly construct a narrative identity throughout living. Therefore, narrative self-making appears to be a lifelong process through discovering our unique traits in social interactions, using patterns of repetition and variation as well as cultural and collective patterns to shape their never-ending evolution.

### **1.7. Towards characterizing proto-narratives in early vocal interaction**

In this chapter, we presented a wide range of research on various key characteristics of narrative. This overview brings us to David Rudrum’s assessment (2006) that narratives are so familiar to us and easy to identify. It is challenging to create a genuine taxonomy to analyze them. This fact could also enable the

identification of exceptional components. The search for defining characteristics may thus benefit from in-depth explorations beyond the question of structure.

The next chapter will present precisely different contexts of early mother-infant interaction and the studies related to the characteristics of early vocal interaction that contribute to forming narratives. The description of early mother-infant interaction will be made in light of different elements and their integrity: first, the intuitive and affective approach by the mother to share and communicate with her baby; second, the willingness of expression by the infant and third, the shaping of shared emotion by both parties. To go over an attempt to answer the previous question again, proto-narratives in early vocal interaction should be considered as a process incorporating narrative-like structures.

Considering any kind of narrative structure elucidated above, proto-narratives in early interaction should be analyzed with many variables included. Interaction and context are the critical markers for drawing the contours of the research. The linguistic asymmetry of the conversational partners makes mother-infant interaction a peculiar form of communication. The mother's intuitive adaptation to her infant, especially to the musical qualities of vocalizations, through minimizing verbal content and maximizing acoustic properties of her discourse, enhances this interaction's rhythmic and melodic dynamics (see Chapter 2 for further details on maternal speech).

Revisiting narrative forms and structures given this particular context is essential to conceptualize proto-narratives in early interaction. However, the complexity of interaction among modalities and dimensions, along with the focal period of life, challenges the process of conceptualization. The following chapters will build in additional views and efforts to this attempt.

## Chapter 2: Contexts of Early Mother-Infant Vocal Interaction

When we investigated how infants begin to communicate, we found less research on preverbal communication and even less on the pre-babbling vocal stage. Most research on early infant communication has focused on how caregivers interact with their young infants in terms of acoustic properties; infant responses have yet to be studied. Hence, we would like to focus on this critical period at the beginning of life and explore the rudimentary nature of communication.

Musical fundamentals of language, such as prosody, are the primary communication features human infants encounter. Research on language acquisition is mainly focused on the babbling stage, which is straightforward in terms of analysis. The pre-babbling stage requires a different approach in terms of analysis. Therefore, researchers have focused on the acoustic qualities of vocalizations for proper categorization.

This chapter reviews research on early mother-infant vocal interaction. Our focus is on the pre-babbling stages of vocal development, when infants' vocal capacity and awareness enable them to produce meaningful vocalizations when engaged in interaction. We present studies examining infants' vocal capacity in the pre-babbling stage, representing their willingness to express themselves, especially in interaction. We also present characteristics of maternal speech during this period based on acoustic properties and verbal content. Studies on how maternal interactive styles are prolific in conjunction with cultural variety and parenting styles will be in focus to discuss the place of musicality in interaction with infants. Individual vocal expressions will also be described through the lens of turn-taking studies, pointing out that micro-interactive units are already present in the early mother-infant interaction. We aim to understand how each interactive unit joins together to form narrative formats on many levels. Each level contributing to storytelling are subject of the research presented in this chapter.

While communicating during the pre-babbling stage, the mother and infant create various joint formats of interaction (Bruner, 1985). These formats appear in many forms of vocal interaction and afford active infant participation. We review various interactive styles, evaluating both speech and musicality, such as singing, as an example of a narrative format used in early interaction. We will further describe the 'proto-narratives of early vocal interaction' based on many different levels of storytelling form in interaction. We examine how these levels or units unify and create a continuum in early vocal interaction.

The focus here is on the musicality of early vocal interaction. Thus, we review every aspect of early mother-infant vocal interaction through Langer's non-discursive frame (1942), putting emphasis on the musicality of individual and interactive vocalizations. Early vocal interaction is beyond discursive and must be studied based on its symbolic function (Langer, 1953) regarding meaning-making units participating in the process of stringing together to become a unity. Symbolic function, according to Langer, does not have a reference because no verbal expression is suitable enough to articulate it. Thus, due to its characteristics, early mother-infant vocal interaction embodies symbolic meaning rather than purely discursive meaning. From this, we will discuss how small vocal units contribute to narrative continuity and how narrative unity begins to be conceived as a simultaneous expressive form in early vocal interaction.

### **2.1. Early Encounters through Sound**

Before birth, the infant is already found to be in interaction through sound exposure in utero. These early encounters arise in various forms, such as speech, storytelling, songs, or nursery rhymes. Sound exposure during pregnancy contains speech characteristics and narrative forms recognized straight after birth.

According to Cheour-Luhtanen et al. (1996), the fetal auditory system is mature enough long before the end of gestation, starting from the 25th week. Recurrent prenatal exposure to a rhyme for four weeks during gestation influences the near-term fetal heart rate, eliciting a decreased response for the familiar versus unfamiliar rhyme (De Casper et al., 1994). Prenatal exposure to the mother's voice affects the

recognition of this familiar sound immediately after birth, and this hearing exposure influences their preferences immediately after birth. Human infants can discriminate their mothers' voices from those of a female stranger's voice (Cooper & Aslin, 1989; De Casper & Fifer, 1980). Newborns' interests are not limited to their mothers' voices. They also respond to familiar nursery rhymes, both sung by the mother or by a female stranger (De Casper & Spence, 1986).

Mehler et al. (1978) showed that infants at one month of age are attentive to prosodic features of maternal speech. When the prosodic aspects of their mother's speech are eliminated, infants' interest fades compared to vocal exposure, including prosodic components. Newborns are also sensitive to specific prosodic features such as linguistic rhythms. A study showed that French newborns successfully discriminate between foreign languages from different rhythm groups but fail to discriminate against foreign languages belonging to the same rhythmic group (Nazzi, Bertoncini & Mehler, 1998).

Both prenatal and newborn studies have demonstrated that the ability to distinguish speech sounds from musical sounds is present. The heart rate responses of near-term fetuses differ according to foreign language speech sounds and a musical melody showing that even near-term fetuses can discriminate between speech prosody and music (Granier-Deferre et al., 2011). Researchers have also performed a study that demonstrated that recurrent prenatal exposure to a descending melody elicits cardiac responses to the same melody when the infants are one month old (Granier-Deferre et al., 2011). This study showed that infants process different sounds and retain recurrent sounds in their auditory memory.

Early sound exposure allows infants to encounter linguistic and sonic forms of expression. These forms are recognizable for infants beginning from early life and impact future symbolic features of feeling, interacting, and thinking.

## **2.2. Post-Birth Vocal Interaction**

During the first year of life, linguistic properties of communication are gradually acquired through interaction via multimodal ways of expression, such as gaze, gesture, bodily movements, vocalizations, and babbling. Research focusing on early vocal development has shown the importance of interactional settings and their impact on shaping infants' vocal production. How mothers and infants express themselves during early conversations are important markers of how the interaction gradually develops into patterns, including specific characteristics similar to storytelling.

### **2.2.1. Infant-Directed Speech**

Parents speak to their infants in various cultures in specific ways with pitch variation, heightened rhythm, repeated melodies, and speech content. Indeed, when engaging with them, parents tell stories to their infants. These stories vary across cultures, including songs, lullabies, lived or fictional stories, and nursery rhymes.

In Western contexts (European and North American middle and upper-middle classes), infants have been perceived as conversational partners since birth. Their vocalizations were taken as responses while primary caregivers talked to them. Mothers naturally engage in face-to-face interactions. American English is the most studied language among Western infant-directed speech (IDS) studies.

Most research on vocal communication between parents and infants has been conducted in European and North American populations (Beebe et al., 1985; Gratier, 2001, 2003; Papousek & Papousek, 1989; Trevarthen, 1977). Researchers have described the expressive vocal interactions that infants partake in actively from the age of approximately two months. These are often referred to as "proto-conversations" (Bateson, 1979; Snow, 1977; Trevarthen, 1974) since they are composed of alternating and simultaneous vocalizations in face-to-face interactions (Stern et al., 1975). Simultaneously, infants smile and look closely at their mothers' faces (Lavelli & Fogel, 2005; Messinger & Fogel, 2007).

Furthermore, many studies have shown that adults spontaneously modify their speech style while addressing their infants. This particular speech was presented first

as the baby talk register by Ferguson (1964) according to research conducted in six different languages, including Syrian Arabic, Spanish, American English, Marathi, Comanche, and Gilyak. Ferguson specifies three major modalities of baby talk: modification of intonation, such as high pitch and overuse of some pitch contours; modification of some words or morphemes and exclusively invented words for babies.

This register, later called *Infant Directed Speech* or *Motherese*, was studied in European and North American contexts (Cooper & Aslin, 1990; Fernald, 1985; Fernald & Simon, 1984; Stern, Spieker & MacCain, 1982), including inter-language comparisons (Fernald et al., 1989) according to which many important similarities have been found between the speech of Europeans (French, Italian, German, English) and North American and Japanese mothers when they talk to their infants. For example, Fernald et al. (1989) reported that French and Italian mothers have a similar range of fundamental frequencies; however, French mothers have higher maximum frequency scores than Italians. Despite minimal variation between cultures, a common high-frequency pattern in the maternal speech was found among the six groups.

Some research on other populations of different cultural contexts shows the existence of similar vocal exchange trends in the rural environments of Southern and Northern India (Keller, 2007; Gratier, 2003; Stork, 1986), urban Nigeria (Trevarthen, 1988) and South Africa (Cowley et al., 2004). Mothers' speech in early vocal interactions seems to present some common patterns in different linguistic and cultural communities. However, peculiar styles have also been reported and are possibly linked to cultural representations (Cowley et al., 2004; Fernald & Morikawa, 1993; Gratier, 2003). As an example, Cowley and colleagues (2004) found that KwaZulu mothers and their 14-week-old babies vocalized simultaneously in a significantly higher ratio. In a study of West African immigrant mothers living in Italy and native Italian mothers with their three-month-old infants, Carra et al. (2013) reported no significant differences in terms of vocal expressions between the two groups, suggesting that the new linguistic environment could influence immigrant mothers' interactive styles. However, West African mothers are found to be more involved in rhythmic vocalizations and singing, which carries the signature of their culture of origin, whereas Italian mothers prefer talking.

Poetic features were also in focus while investigating babytalk. Miall and Dissanayake (2003) analyzed a one-minute vocal interaction of a Scottish mother and her 8-week-old baby. Authors focused on the poetics of the babytalk through a detailed analysis of poetic features of maternal speech. First of all, they underlined temporal frames defined by precise beginning and endings, along with repetitive metric features such as pauses and stresses and utterances organized in lines and around a theme, announcing the macro-poetic structure of the babytalk (2003: 343). Second, authors precisely investigated micro features like verse patterns and metrical structure within the lines. This analysis also reflected elements of poetic foregrounding, highlighting a shift of emotion or attention for recalling baby's engagement in the interaction. In this example, foregrounding effect is caused by phonetic elements, precisely by stressed syllables. Authors claimed that an early responsiveness to aesthetic features of babytalk is present, enhancing the infant's future emotional and cognitive abilities.

#### **2.2.1.1. Melodic Properties**

Maternal speech prosody is a significant aspect that many researchers studying speech directed to infants have focused on. Infant-directed speech (IDS) has been distinguished from adult-to-adult speech (Garnica, 1977; Papousek, Papousek & Bornstein, 1985; Fernald et al., 1989) mainly by exaggerated pitch contours, increased use of high pitch, repeated melodic contours (Fernald & Simon, 1984; Papousek & Papousek, 1989; Gratier & Devouche, 2011), important variations in fundamental frequency (Garnica, 1977), longer pauses (Stern et al., 1977), and lengthening of vowels (Fernald et al., 1989).

It has been reported that babies prefer IDS from the first month of life (Cooper & Aslin, 1990), at four months; infants show a significant interest in motherese even spoken by strangers rather than adult-directed speech (Fernald, 1985). Among these acoustic properties, infants show a significant preference for frequency patterns of IDS rather than other properties such as amplitude and duration (Fernald & Kuhl, 1987).

Scherer (1979, cited in Fernald & Simon, 1984) showed that high pitch is related to positive affect in adult-to-adult speech. Fernald and Simon (1984) also noted that



rising contours tend to arouse the baby, and falling contours are related to calming the baby down (Papousek et al., 1985). IDS has also been suggested to have affective qualities (Fernald, 1989; Kitamura & Burnham, 2003). Smith and Trainor (2008) tested high-pitch use in an experimental strategy where the mothers believed their babies could hear them from another room. However, an experimenter controlled the babies' responses by positively reinforcing high-pitch use in maternal speech. Researchers have found that the rise in pitch is highly influenced by the positive feedback given by the infant.

Infant responses are an essential aspect of analysis because they influence mothers' behaviors, and as the baby grows up, the characteristics of IDS change synchronously (Stern et al., 1983). Therefore, mothers' performance and style are highly affected by infants' reactions.

#### **2.2.1.2. Signature tunes**

In addition to the specific acoustic properties of maternal speech, there is a further recognizable feature in prosody and music called signature tunes. This feature may contain particular melodic or rhythmic phrasing distinguished by its repetitive use. It also gives an identity to a peculiar sound pattern that the person possesses, as Lewis describes as "having your own sound" for jazz musicians (1996:117), and this makes the musician "someone" by assigning a specific sound identity (Iyer, 2004).

In Western contexts, a few further studies report that each mother owns a repertoire of specific melodic sequences, "signature tunes," that the infant associates with their mother (Bergesson & Trehub, 2007; Gratier, 2001). Bergesson and Trehub's study shows that mothers frequently use a specific melodic sequence within and across recording sessions while speaking to their infants. Researchers have suggested that these signature tunes are recognized by infants and impact their vocal production.

#### **2.2.1.3. Verbal Content**

Many researchers showed that infant-directed speech also has significant variations in terms of lexical and syntactic structures. It is suggested that simplified structures and inputs help babies access language and facilitate language

development. This aspect remains a subject of discussion by many researchers (Soderstrom, 2007). Several contradictory results have been shown in studies among indigenous societies concerning the differences in parenting styles and cultural strategies (e.g., Ochs & Schieffelin, 1994). Jakobson (1960) interprets the attempt to simplify the speech by parents as a way to find a “common code” between both sides, the one who ideally has access to language and the other who has not yet. Researchers report that concerning the verbal content of the speech, regarding lexicon, mothers use more diminutives and reduplicated words. They tend to make simple and short sentences and use interrogatives regarding grammatical structures. Several properties have been listed to describe the verbal content that mothers address infants.

As we look at how mothers speak to their infants regarding verbal content, repetition appears as a primary result of many studies (Kaye, 1980; Kruper & Uzgiris, 1987; Stern et al., 1983; Papousek et al., 1987). Snow (1972) shows that repetition of words and phrases frequently appear in the mother’s speech. Ferguson (1964) observed the reduplication of words and frequent use of diminutives. Many studies have reported short sentences, such as by Papousek, Papousek, and Haekel (1987), who showed that utterances used in speech to infants aged three months are shorter than those of infants at 12 months of age. Large numbers of questions have been addressed to young infants (Kruper & Uzgiris, 1987; Soderstrom et al., 2008; Toda, Fogel, & Kawai, 1990), both from mothers and fathers (Kruper & Uzgiris, 1987).

Many studies have shown that the theme changes over time in terms of the maternal discourse. In a study of infants from 3 to 18 months of age, Snow (1977) documents that there is a significant change in the verbal content: at three months, the mother’s focus is on the baby, and this echoes in her speech content; however, at 18 months, the conversation is more about the social or physical environment. Pêcheux et al. (1993) observed babies in change situations and reported that maternal speech is centered on the body or the current situation at three months but changes at nine months of age.

Not only age but the baby’s responses also influence this revision of speech content, as we encounter in the example of acoustic properties. Henning, Striano, and

Lieven (2005) studied infants in their first and third months and observed a variation of complexity in syntactic structure and an increase in the number of utterances used in interaction. They also reported semantic variations between these two age groups, which are related to the progressive changes in positive responses given by the infant.

### **2.2.2. Infant-directed singing**

Another widespread register used when mothers address their infants is singing songs; lullabies are primarily found to be universal (Trehub & Trainor, 1998). To put them to sleep or simply to entertain them, mothers sing to their babies, but also toddlers and children sing to their baby dolls (Trehub, Trainer & Unyk, 1993). Singing is a narrative form of expression, holding a structure similar to routines described by Bruner (1990).

Singing to infants is predominant across different cultures; these repertoires share common musical characteristics (Trehub, Unyk & Trainer, 1993), such as repetitive units and the simplicity of their structure (Unyk et al., 1992). From two months of age, infants are able to vocalize, maintaining the musical structure of the song sung by their caregivers (Mazokopaki & Kugiumutzakis, 2009). Singing play songs to infants display the use of high vocal pitch and slow tempo (Trainor et al., 1997; Trehub & Trainor, 1998). Lullabies and play songs show differences in pitch variability; play songs appear more variable than lullabies (Trainor et al., 1997).

Singing directed toward infants and in the absence of infants also differs with regard to the use of high pitch (Trainor et al., 1997); high pitch variability was found in infant-directed versions of play songs. However, no differences were found between the two conditions for the lullabies. In a similar comparison, Nakata and Trehub (2011) found that infant-directed singing has more salient temporal regularity than non-infant-directed singing.

Infant-directed singing also shares common qualities with infant-directed speech, such as louder or slow singing when adults want to arouse the baby's interest and longer pauses between utterances (Trainer et al., 1997).

Repetitive patterns in lullabies and children's songs are a way to introduce them to ritualistic patterns (Merker, 2009). Like singing patterns, peek-a-boo play also appears as a main conversational style among different cultures (Bruner & Sherwood, 1976; Greenfield, 1972) and is a critical way to access culture and language through rituals.

Various infant-directed expressions own crucial characteristics that induce infant participation. Mothers speak to their infants verbally and musically. This particular speech possesses narrative elements such as verbal stories, songs, and repetitive patterns that grant the infants to enter the realm of creating and sharing stories.

### **2.2.3. Infant's Vocal Repertoire**

Understanding infants' vocal capacity in early life is vital to recognize their willingness to participate in a conversation. Infants' active participation is the major thread of forming meaningful interactive sequences and building narrative formats essential to their language development.

Immediately after birth, infants begin to produce sounds. The cry is the primary way of expression. Natural exposure to the maternal voice in the womb dramatically impacts an infant's vocal production at birth. Research has shown that French and German newborns can reproduce the prosodic components of their respective maternal tongues' intonational contours in their cry tunes (Mampe et al., 2009). Most research on infant vocalization excludes non-speech-like vocalizations due to the difficulty of the study. In a recent study, Yoo, Bowman, and Oller (2018) highlight the importance of mothers' reactions to their infants' cries. Researchers here preferred to include cries, arguing that the response to crying is also an important aspect to take into account to understand the nature of the contingency of the relationship between mothers and infants.

Infant vocalizations during the first weeks of life have vital qualities. These early vocalizations are called vegetative sounds associated with biological needs and reflexive for those expressing distress. General vocal categories were also admitted,

such as speech-like (protophones) versus non-speech-like or cry versus non-cry from birth up to approximately the second month. Many researchers who worked on the classification of vocalizations during the first year of life present slight chronological differences in the borders of the phases of the pre-babbling stage.

Oller (1980) defined the vocal production of infants during the first year of life in his classification system for the emergence of speech sounds in infancy. According to him, infants usually produce guttural sounds of a vegetative nature in the pre-canonical period, from birth to the end of the first month, which he calls the phonation stage. From 2-to-3 months old, infants vocalize velar consonants during the goo stage, which are typically non-reduplicated and unique. Oller (2000) described the sounds infants produce in the early pre-babbling stage as pre-canonical. These vocalizations are different from those he describes as canonical produced during the babbling stage, accepted as the indicator that the baby gets into the linguistic field (Oller, 1986; Oller & Lynch, 1992).

Stark (1980) extends the first stage of vocal development until the end of the second month and illustrates infants' vocalizations during the first two months as reflexive with vegetative sounds. The second stage, corresponding from the third to the fifth month, is the cooing stage and includes the first voicing and laughter.

As stated by many, the second month is pivotal in terms of the vocal qualities of infants as well as the acknowledgment of this quality by caregivers. According to Stark and colleagues (1993), the 2-to-3-month juncture includes not only new developments in vocal capacity but also carries old patterns from an early stage. Stark claims to include non-speech-like sounds in the classification, suggesting that these sounds are fundamentally crucial for the development of speech through a regression-progression process (Stark, 1980).

Koopmans-van Beinum and Van der Stelt (1979) considered an infant's anatomical capacities based on phonation and articulation to describe vocalizations of the very early stage, suggesting that respiration impacts the quality nature of vocalizations. During the first two months of life, they reported uninterrupted

phonation. From the second until the end of the fifth month, they described the vocalizations of this period as interrupted phonation with one articulatory movement.

Following the importance of his work to the existing literature on infant vocal production, research on volubility is also part of Oller's work. Oller and his team define volubility as the "amount of talk" and reported that less research was dedicated to infant volubility rather than parental one (Iyer et al., 2016; Oller et al., 1994). Studies on interactive and non-interactive situations before the babbling stage have shown that infants are more engaged in interactional social situations. Besides, they significantly contribute to reengaging their mothers in conversations (Franklin et al., 2014). In other words, an increase in vocalization in a non-interactive situation is interpreted as a call for participation from the mother. This transition from the pre-canonical to the canonical stage is also a milestone for infants to enter the world of conventional linguistic rules.

Bateson (1975) reported that maternal vocalizations influence the frequency of infant vocalizations from the second month of life in eye-to-eye contact situations. Following this model, Bloom and colleagues conducted studies to describe infants' vocal qualities during the conversation. Individuals listening to recordings of infants with their conversational partners describe that infant responses give them the impression that they "really talk" (Bloom, Russell & Wassenberg, 1987). Mothers' responses also appear to be contingent on these vocalizations (Masataka & Bloom, 1994). Babies modify their vocalizations acoustically and temporally in the way that their mothers address them. When alone, they produce short-timed and less-modulated vocalizations (Bloom, 1988).

An infant produces meaningful sounds along with vegetative sounds from a very early stage when interacting with a conversational partner. Nevertheless, we observe more qualities of their skills of communication. The vocalizations of early age that we prefer to describe as *music-like*, in turn-taking, or synchrony with caregivers, have been the subject of many studies in terms of rhythmic and temporal organization (Gratier, 2003; Jaffe et al., 2001; Stern et al., 1977; Stern et al., 1975; Trevarthen, 1977). Studies have shown that vocal interactions are organized around similar

temporal intervals and have high rhythmic coordination between mother-infant couples. Rhythm and synchrony appear to be fundamental expressive components of early infant behavior in interactional settings, influencing future behavioral outcomes (Jaffe et al., 2001). Thus, rhythmic coordination and interactive timing experiences can also be considered a rudimentary basis for self-expression. In research conducted on musical stimulus and motor coordination, Zentner and Eerola (2010) reported that infants from 5<sup>th</sup> to 24<sup>th</sup> month-old are more involved in rhythmic behaviors when they listen to music compared to the rhythm of speech and the rhythmic coordination of music and movements are associated with the exposition of positive emotion.

However, few studies have examined the characteristics of melodic contours produced by infants through vocal matching analysis (Gratier & Devouche, 2011; Kokkinaki & Kugiumutzakis, 2000; Papousek & Papousek, 1981, 1989a). Research on vocal imitation shows that infants at three months old can precisely reproduce the contour and pitch of their vocal partners (Kessen, Levine & Wendrich, 1979; Kugiumutzakis, 1985). Papousek & Papousek (1981) discussed that rhythm precedes melody in the vocal productions of infants, regarding its multimodal biological roots, such as motor behaviors. Later, these authors studied the vocal matching of melodic contours between mother-infant dyads (1989a). A significant increase in the number of infant vocalizations was found between two to three-month-old, confirming previous findings (e.g., Stark et al., 1993). The majority of matched melodic contours were observed in fall and bell-shaped forms, in parallel with their previous study in 1981. A comparable research on vocal imitation in the third month reported that infant vocalizations were also found similarly in the form of bell-shaped or ascendant/descendant contours among pairs of successive vocalizations (Gratier & Devouche, 2011).

An infant's vocal capacity and musicality of vocalizations are important indicators for examining their participative presence while shaping the organization of the interaction. It is essential to recognize their role during the formation of communication.

#### **2.2.4. Turn-Taking**

Some specific ways to communicate with infants and infants' vocal participation were the subject of multiple studies. Further questions have also been addressed to study how successive vocalizations are combined to form specific interactive units during conversations between mother and infant. These modules, called turn-taking sequences, are markers of conversational continuity and take part in building levels of narrative unity.

Sacks, Schegloff & Jefferson (1974) defined taking turns during a conversation through two modalities: turn constructional units, short or long utterances allowing possible prediction of the end and transition of relevant places, and potential gaps to take the turn. Many researchers have demonstrated the existence of turn-taking registers in adult-to-adult conversations (Sacks et al., 1974; Sidnell & Stivers, 2012), as well as in mother-infant interaction (Gratier et al., 2015; Levinson, 2016).

Turn-taking in mother-infant interaction is formed with the mother's discourse and takes shape with spontaneous vocal responses of the infant; one altering between these vocalizations is needed to identify a turn-taking unit. These units comprise separated infant and maternal vocalizations and overlapped vocalizations (Bateson, 1975; Bloom, 1988; Gratier & Devouche, 2011; Gratier et al., 2015; Stern et al., 1977; Trevarthen, 1977).

The duration of the expressive and silent units is an important marker of perception and contingency. Many studies have shown that taking turns between a mother and an infant of 2-to-4 month-old last shorter than one second (Beebe et al., 1985; Keller et al., 1999; Stern et al., 1977; Stern & Gibbon, 1979; Weinberg & Tronick, 1994). Predominantly, it is reported that pauses between vocalizations vary from 500 ms to 1 s (Jaffe et al., 2001). These silent intervals rarely last more than 3 seconds; in this case, they are considered an interruption or a broken rhythm of the social engagement (Stern et al., 1977; Stern & Gibbon, 1979). In infant vocal studies literature, 3-second intervals are considered a marker of social contingency (Stern et al., 1977; Van Egeren, Barratt, & Roach, 2001). This specific temporal harmony influences an infant's vocal quality (Bloom, Russel & Wassenberg, 1987) and is



appreciated by the partner as intentional and expressing emotions (Papousek & Papousek, 1989b).

Various forms of vocalizations within turn-taking units have been studied. Overlapped and latched vocalizations seem to be more interesting. Gratier et al. (2015) studied turn-taking units of 2-to-5-month-old infants of French dyads. Of infant vocalizations, 44.5% were latched turns, an infant vocalization following a maternal vocalization within less than 50 ms. A high frequency of latched vocalizations may represent an infant's anticipation of taking turns. Still, the number of overlapped vocalizations is higher than that of latched vocalizations. Depending on who initiates the overlapped vocal unit, latched and overlapped turns should be reconsidered to analyze the anticipation hypothesis during this pre-linguistic stage. Overlapped vocalizations are also very interesting to work on, while it is unattainable to analyze them properly because of the lack of technical support separating the two sounds. To properly analyze simultaneous and separated vocalizations within overlapped units, the recording procedure should be performed in isolated rooms for the mother and infant to obtain convenient acoustic measures.

Like many psycholinguistic studies, we also encounter cultural-specific differences in taking turns in a conversation, which is a pattern commonly accepted as universal (Brown, 1998). However, recent works in this field, including multiple comparisons of 10 languages, provide some output on similarities between cultures (Stivers et al., 2009) through approximately 250 ms of pause duration for answers to questions, although reporting minor variations across cultures.

Besides turn-taking units, phrasing is an essential property of music and speech. Lerdahl & Jackendoff (1983) described that musical motifs are ingrained hierarchically to build musical phrases, and multiple phrases form themes and, finally, a complete piece of music. At the same wavelength, Lynch and colleagues (1995) examined acoustic and temporal parameters of infant vocalizations from 2 to 12 months of age using free evaluators. They reported that infant expressions composed of vocalizations and pause breaks are organized in phrases. Phrase-final syllables were more extended than non-phrase final syllables within phrases for 2-to-4 month-old infants.

Reciprocally, in research on how infants perceive the structure of musical phrases, Jusczyk & Krumhansl (1993) reveal that parsing is marked by indices such as lengthening the end with a fall of fundamental frequency ( $f_0$ ).

Turn-taking sequences designate the example of micro-units combined to form extensive interactive formats and the infant's part throughout this joint formation. These sequences already display that infants are closely capable of following communicative patterns.

### **2.3. Early Socialization Practices**

As many studies referring to the features of speech directed to infants' states similar characteristics, research broadcasting cultural variations suggest diverse interactive styles and ways of communicating with the infant. Cultural-specific norms influence how some communities use interactive strategies and present a variety of different forms of communication. Thus, the existence of common strategies is also presented in light of language socialization studies. This section presents a diverse perspective on speech, prosody, and narrative practices considering cultural-specific patterns.

#### **2.3.1. Language Socialization**

Socialization is the process of accessing the norms, values, and ideologies of a social group, community, and culture. Communication, language, and cultural values were blended in terms of information for the researcher while examining the vocal development of infants. The socialization process starts long before language acquisition; from the first moment, the infant becomes interactive, even if the caregiver does not give the infant this status.

Many linguistic studies have been conducted among Mayan communities in Mexico on the language acquisition and language socialization processes of infants and toddlers (Brown, 1998; De Leon, 1998; Gaskins, 1996) and also among other indigenous communities such as Gusii and Kaluli (LeVine et al., 1996; Schieffelin, 1990). These observations suggest that the baby progressively becomes an active actor in

conversation and contributes to social interaction with close family members through the use of linguistic structures.

The status of the infant in the pre-speech interaction stage and communication styles, depending on the participants, vary across cultures. In European contexts, the primary mode of communication is dyadic face-to-face interaction (e.g., Fernald et al., 1989); however, different modes of communication have been described, such as triadic or multiparty interactions among other communities. In this type of interaction, an infant has the status of an observer and access to language exposure through overhearing or reported speech (De Léon, 2011; Ochs & Schieffelin, 1994). Young infants are not labelled as conversational partners, as we observe in European social contexts (Brown, 1998). In these communities, infants were allowed a spaced-out observation scene with more salient multimodal (verbal and nonverbal) expressions from all participants. For example, in Africa, infants have more than one conversational partner (Kärtner et al., 2010; Keller et al., 2005; Rabain-Jamin & Sabeau-Jouannet, 1997).

It is also stated that in Mayan communities, infants' vocabulary consists of verb roots in Tzetzal (Brown, 1998) and Tzotzil (De Léon, 1999) rather than simple nouns, as we mainly observe in European contexts (Clark, 1993).

A significant aspect was broadcasted by developmental psychologists working predominantly in European and North American contexts in which mothers simplify their speech grammatically while addressing their babies (Ferguson, 1977; Fernald et al., 1989). However, this specific aspect of baby talk has yet to be observed among indigenous communities (Ochs & Schieffelin, 1994). Another critical aspect of infant-directed speech, repetition for stressing new information, has also been observed in Mayan communities (Brown, 1998). However, in this case, it appears as a language-specific component rather than an act simplifying language acquisition.

Despite the differences observed at the syntactic and grammatical levels of speech for infants, and multiparty versus face-to-face dyadic interactions, one aspect appears common in both viewpoints, and that is on the basis of modulation of acoustic parameters of the speech:

“Language addressed to the young infant tends to be in the form of songs or rhythmic vocalizations in a soft, high pitch.” A Samoan Developmental Story in Language Acquisition and Socialization: Three Developmental Stories and Their Implications (Ochs & Schieffelin, 1994, p. 489).

Brown (1998) noted that speech addressed to Tzeltal babies is often whispered. Ochs & Schieffelin (1994) reported that Kaluli mothers do not consider their infants as conversational partners; however, in a multiparty conversation, when older children talk to them, mothers respond on behalf of their infants with a “high-pitched nasalized voice” (p.483). In another example from the same work, researchers report that Samoan mothers change their voice, becoming “loud and sharp” while the speech is not simplified lexically or grammatically (p.489). In this research comparing three different developmental stories, including a White-American middle-class family, researchers showed that the baby talk register had not been observed among indigenous families, as the universality of IDS suggested by developmental psychologists working mostly in North American and European contexts. However, some prosodic changes were observed even though a difference existed between registers to address the infant.

In Western contexts, the status of an infant is considered a conversational partner from birth but treated as a foreigner who does not master the language solely like a native speaker (Ferguson, 1977; Snow, van Eeden, & Muysken, 1981). Thus, most research in Western contexts reports that a syntactically and lexically simplified register is used while addressing the infant. Otherwise, in non-Western communities, an infant is not taken as a conversational partner in the early stages of life; thus, mothers do not simplify their speech to avoid restraining the linguistic exposure of the infant, according to an example from the Kaluli community (Ochs & Schieffelin, 1994).

The existence of a generalized register of talking to babies should be reconsidered in light of ethnographic and intercultural studies, and acknowledging the culture-specific patterns of parenting and socialization is necessary. Musicality, however, is entitled to be acknowledged as increasingly present in research addressing infant communication.

### **2.3.2. Parenting Styles**

Our consideration of how cultural patterns affect socialization and language is similar to the aspects reported by Ochs & Schieffelin (1994), Schieffelin & Ochs (1998), and Greenfield, Brazelton & Childs (1989); cultural patterns should not be acknowledged independently for these values which are blended into communication styles along with language usage and are grounded in our communicational development. Along with Cowley, Moodley & Fiori-Cowley (2004), we argue that language is not just an output but also a fundamental transmitter of culture. Alongside these aspects, socialization appears as a multimodal developmental process through different modes of interaction.

Parenting styles, in other words, how primary caregivers are engaged interactively with their babies, show diversity between cultures. Parenting impacts an infant's development from birth by encountering cultural and social environments (Greenfield et al., 2003; Keller et al., 2004; Keller, 2003; LeVine, 1970).

Papousek & Papousek (1989b) described "intuitive parenting" in Western contexts, suggesting that all parental behaviors are adapted spontaneously to their infants. While engaging in conversations, parents adjust to the responses of their infants in terms of intensity and quality (1989b, 2002). As for vocal development, acoustic alterations in parents' discourse favour the infant's apparent perception of linguistic units (1989b). Papousek (1986) also mentioned that musical games, singing, and infant-directed speech are the earliest forms of music education for infants with no intention of teaching the baby proper musical knowledge but to share time and enjoy it mutually.

Greenfield, Brazelton, and Childs (1989) reported that Zinacanteco (Mayan) mothers hold their babies close to their bodies until the end of the first year of life, which is quite the opposite of what has been observed among mothers in the United States. Haviland (1978) mentioned that Zinacanteco mothers find that this position is an adequate form to keep the baby comforted.

Studies comparing Western and non-Western populations on the characteristics of parenting styles have also been conducted. Keller, Voelker & Yovsi (2005) documented in a comparative study between German and Cameroonian Nso mothers that body contact and stimulation are more commonly observed among Nso mothers than among Germans.

Keller (2007) later describes two approaches to parenting styles: proximal and distal. Keller et al. (2009) report that proximal styles such as face-to-face interaction and object stimulation are commonly seen among urban well-educated Western communities. However, in low-educated non-Western rural contexts, proximal parenting behaviors such as excessive body contact have been inspected. Mutual gaze is also one of the important aspects of the distal pattern of early engagement. Parents using the distal strategy are more in mutual eye contact with their infants (Lavelli & Fogel, 2002). Carra et al. (2013) studied mother-infant dyads in the context of immigration. First-generation immigrant West African and Italian mothers were in focus. This research documents, as expected, that West African mothers show a more proximal style of parenting than Italian mothers.

Further cross-cultural studies observing infants up to six months report variations between cultures. Bornstein et al. (2012) conducted a naturalistic observational study in rural and urban settings in Italy, Argentina, and the United States to monitor emotional relationships between mother-infant dyads. Results suggested that Italian infants are more emotionally responsive, and mothers are more receptive to their infants compared to their U.S. and Argentinian counterparts. In further research conducted among eleven countries, including France and Italy, Bornstein et al. (2015) report that common contingent vocal sensitivities were found among cultures regardless of differences in vocal rates. No correlations were found between maternal and infant vocal rates, except for Brazilian and Argentinian dyads, a positive correlation is noted between rates. However, mothers and infants were highly responsive to each other among all cultures. On a responsiveness comparison between interaction partners, Italian, Belgian, and U.S. American mothers are found to be more responsive than their infants, while other communities are described as equally

responsive dyads. No significant difference was found in infant vocal rates across cultures.

It is also reported that using both strategies has been detected among urban-educated families. The alternation of choosing one of these strategies is related to different intentions. The effects on the socialization of the infant at three months of age have been shown in this cross-cultural work by comparing other socio-demographic groups from different cultures as well as within the same culture. Distal parenting strategies are recognized being a way to support early awareness of self (Keller et al., 2005; Keller et al., 2004).

In a study on compared vocal interactions, Demuth (2008) worked on Nso and German infants at three months and their mothers. This study reports that both interactions of these communities present musical and rhythmical patterns called protosongs (Cowley et al., 2004). However, Nso dyads have shown more simultaneity along with repetitive rhythmic patterns. Besides, further corporal movements were observed as a way of communication among Africans rather than Germans. In Western communities, it is also observed that more overlapped vocalizations are produced in mother-infant interaction than in adult-to-adult dialogue but still less than in non-Western communities. Similar results were found in Indian communities (Gratier, 2001; Stork, 1986), reporting that interactions observed in non-Western communities are more based on body contact than in Western communities, primarily based on vocal communication.

We come across interesting research on parenting styles and parents' narratives to teach their children how to overcome their anger, as realized by Anthropologist Jean L. Briggs. In 1963, Briggs conducted a field study among the Inuit community living in North Western Canada to observe emotional patterns in everyday life context. Briggs (1970) published her work in her book "Never in Anger," a portrait of an Eskimo family. Briggs noticed that this community does not express anger in a typical way. Instead, Inuits see expressing anger as childish behavior and a sign of weakness. When their children make a mistake, mothers in this community do not punish or yell at them. Yelling is equivalent to humiliating them. They claim that children sometimes behave

to make others angry, but it is important to try to understand what makes them upset instead of getting angry.

To teach their children how to overcome their anger, parents tell them narratives. These narratives are created to shape the child's immediate behavior according to a specific situation instead of giving them implicit moral lessons. For example, suppose a mother wants to keep her children away from the ocean instead of giving them an order not to go too close to the water. In that case, she tells a story about a monster that kidnapped children who got too close to the water and put them in his bag to give them to another family. Grownups interpret these stories as funny. Another story example reports that to make the children put on a hat when they go out, and parents say that if they do not wear a hat, northern lights pick up their head and use it like a football to play.

Telling stories is seen as a universal pattern of humanity; various forms are applied in diverse cultures to communicate with them or educate babies and children to prepare them for life. Cultural-specific parenting styles significantly impact using different story forms to address infants and children.

#### **2.4. Towards conceptualizing narratives in early mother-infant vocal interaction**

This chapter covered different characteristics of infant-directed maternal speech, its impact on infant vocal production, along with cultural differences affecting the relationship and other forms of interaction. Musical features of the interaction were introduced to describe this particular interaction as more music-like than merely discursive. Some musical and linguistic subunits, such as acoustic properties and turn-taking units, were presented to consider later how they are unified to form broader units that we will examine in light of various narrative forms. Additionally, infant-directed songs were presented to introduce the idea that narrative forms are already present in early interaction. These studies are essential to comprehend the process of forming expressive units already manifested in an infant's early periods.



Both linguistic and musical units are intertwined while describing this particular interaction. The verbal content of maternal speech differs from adult-to-adult content. However, most importantly, the speech addressed to infants is distinct for its musicality. This musicality component positions early interaction into a symbolic realm, conveying various forms and functions transcending language. This fact leads many researchers to quest new ways of meaning-making within the bounds of interaction.

The next chapter will emphasize musical features of early interaction by bringing musicality into focus and further describe musical narratives in mother-infant interaction. We also discuss how narratives are defined in musical work and while creating music.

## Chapter 3: “Communicative Musicality”: We Are All Born Musicians

After silence, that which comes nearest  
to expressing the inexpressible is music.

Aldous Huxley, *Music at Night and Other Essays*, 1931.

Music is defined as a “humanly organized sound” by Blacking (1973). Musicality, however, is defined as a feature of basic cognitive and biological mechanisms and their functions through the developmental path (Honing et al., 2015). Research on the universals of musicality has shed light on the evolution of language (Fitch, 2010). As Bruno Nettle (1983) says, “all societies have vocal music” despite the absence of “absolute” universals (Savage et al., 2015).

This chapter is dedicated to music and musicality in light of developmental, music, and neuroscientific research. Interactional settings, neurobiological networks, and key concepts are discussed through these research frameworks. Initially, narratives and narrative components will be held as part of human musicality through language, and will then be discussed considering purely musical forms. Some music research will also enhance our understanding of the similarities between musical and linguistic expression.

### 3.1. Communicative Musicality in Mother-Infant Interaction

Until recently, attention to very early mother-infant vocal interactions was missing in the developmental and linguistic research field. However, a newborn’s first encounter with the other is attained through sound. Ullal-Gupta et al. (2013) suggest that the sonic environment in utero has a significant influence on shaping the basis of melodic and rhythmic motifs of one’s culture through musicality and speech prosody.

For the period following birth, Papousek & Papousek (1981) describe baby talk as the first musical encounter of an infant through language. This study has drawn

attention to the musical nature of infants' interactions, early musical capacities, and parents' intuitive responses.

Communicative Musicality (Malloch, 1999; Malloch & Trevarthen, 2009; Trevarthen, 1999) in mother-infant interaction comes into sight as an innovative concept at the very beginning of the 21<sup>st</sup> century. Communication here represents, on one side, literal verbal expression from one partner; on the other, it also contains nonverbal expressions from both interlocutors. Meaning appears as a unifying element of shared interaction reflected on both partners. In other words, meaning has to be considered by recognizing the present moment of the interaction. An infant's contribution gives this interaction an improvisational nature and has a pivotal effect. An infant is not only a conversational partner but, above all, is the protagonist of this particular interaction who has a profound impact on creating the conversation. The musicality of the interaction consists of the prosody of maternal speech and vocalizations of the infant, which are combined and create the entire melodic contour of their conversation. The melodic contour appears as a result of linking the vocalizations of both partners by taking turns and overlapping vocalizations throughout the interaction. The progression of the melodic contour has a fluctuated nature, similar to an improvisation between musicians. Compared to music, speech intonation does not have a straight emotion-arousing nature. However, the mother's specific use of intonation has driven attention and distinguished it from regular speech due to its musicality.

Three crucial elements appear in the theory of Communicative Musicality in mother-infant interaction, and these elements are not restricted to vocal expression (Malloch, 1999):

(1) **Pulse** unifies many elements related to timing, in other words, temporal patterns of communication. Pulse units are marked by the beginning and an ending; these markers allow us to examine the dynamics within the unit. Pulse units are minimal sequential units associated with each other to create a whole unit later that we can call a musical phrase. The duration of these units and their flowing beat are counted while analyzing them.

(2) **Quality** represents the melodic structure of the interaction. The pitch and the fundamental frequency are indicators to define the form of the melodic contour besides the fluctuations and the internal pitch variation. Timber is another dimension of quality, the particularity that makes the sound peculiar.

(3) The third dimension, **narrative**, is described by Malloch (1999) as the combination of pulse and quality. Narratives are essential communication formats of everyday life (Bruner, 1990) through expressing emotions and generating meaning conjointly (Stern, 1993; Trevarthen, 1998).

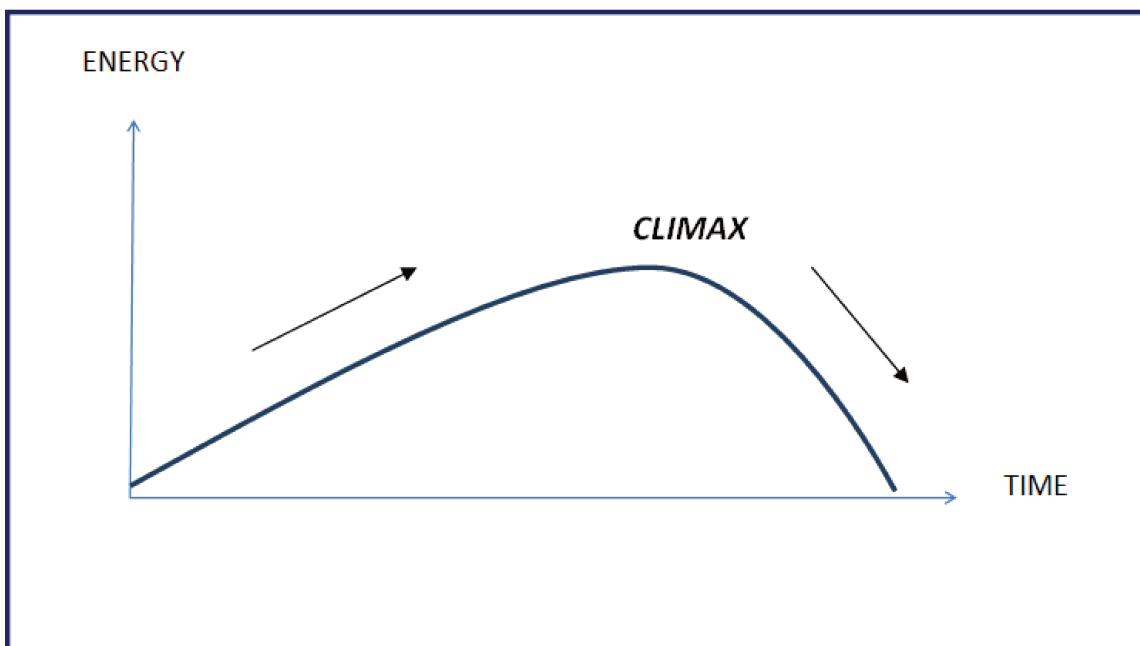
In the first paper on Communicative Musicality, Malloch (1999) presented a narrative form through a nursery rhyme that the mother sings to her baby. This moment between mother and infant involves the nursery rhyme as a narrative, with its lyrics and melody. In addition, the infant's participation through vocalizations makes this narrative unique and musically communicative. Mother jazzes up this uniqueness by modifying her voice or lyrics, creating an exclusive version of the nursery rhyme with her baby. Malloch points out the similarities between the build-up of the units of this nursery rhyme and a regular conversation between mother and infant: the conversation begins and amplifies gradually, reaches a maximal point, and then falls out. Escalated tension and relief of this energy are seen among most musical pieces and create a narrative effect (Imberty, 2005).

The narrative is not limited to voice or language. Corporal expressions accompanying the speech are also exclusive sources of narrative. Body movements trace a narrative-like path and are synchronized with vocal activity. Delafield-Butt and Trevarthen (2015) show that a premature newborn demonstrates synchronized vocal and arm movements in an interactional situation with the mother. These are the kinetic interpretations that the musicality of interaction affords. The vocal climax of the narrative interaction corresponds to raised arm movements in terms of matching intensity and frequency. The climax is described as a "unit of meaning-making" (p.3); i.e., corporal movements co-articulated with vocalizations contribute to the meaning. The infant is fully engaged in the conversation with all possible manners of expression, such as smiles, vocalizations, and gestures. In a previous study, researchers found

cross-modal signatures in speech and singing (Trehub et al., 2013). The experimental setting is based on speech, singing, and silent video clips screened to adults and infants to understand if they could identify previously heard speakers in silent videos. Results confirm the ability to notice cross-modal units by both adults and infants. This study suggested that cross-modality of interaction can be conceived in parallel with the ability of infants to connect remote experiences of perceptual functioning, described by Stern as amodal gestalt (1985/2000). This process of relating isolated experiences allow infants to recognize distinct patterns of the same experience, later coming together to provide the basis for the emergence of the self. Those transfers between different modalities, so their abstract representations tend to merge and become a global unity (Stern, 1985/2000).

**Figure 1**

*Evolution of the narrative dynamic through time*



As seen in Figure 1, the narrative format has this rudimentary pattern of escalation and fall of energy, expressed in multimodal ways. Bruner & Feldman (1993) described the narrative format in a four-stage model: a canonical steady state as an

introduction, followed by a precipitating event, restoration of the event, and a coda for the ending. The authors also suggested that this format is relevant to any kind of interactional setting from the very beginning of life (e.g., peek-a-boo play) for both verbal and nonverbal playful communication. Gratier & Trevarthen (2008) describe that nonverbal narratives are engraved within our everyday verbal narratives. Telling a story is a multimodal activity; involving the whole body to engage partners and captivate and maintain their attention. Similar research on daily routine activities such as nappy changing and book presenting is also studied and reported vocal and behavioral changes in comparable interactional formats (Rossmanith, 2017, cited in Raczaszek-Leonardi et al., 2019; Rossmanith et al., 2014).

Most recent work on narrative formats of social coordination in the early period of life named these formats “action arcs” (Raczaszek-Leonardi et al., 2019). It defined them as participatory patterns being transformed during the social engagement. Researchers describe the existence of such forms with beginning, build-up, and ending contours saying that they afford some future linguistic frames leading the way to more complex further interactions.

In their latest article on the co-regulatory function of narratives, McGowan and Delafield-Butt (2022) propose a developmental framework to review narratives as a unit of meaning-making to understand human cognition and regulatory processes during early development. Authors claim narratives as crucial structures forming interactions and their extension on regulatory and intellectual functions. Each and every embodied experience forms a plot starting from the early periods of life toward a prospective structural basis for knowledge and regulation, similar to the developmental progress of human infants passing through successive growth stages (p.2).

Many researchers have long studied organizational elements within human interaction, which are still ongoing. Insights from early works such as Condon (1976) claim that the organization is a fundamental component of the interaction, just as how we perceive and make sense of the world. Condon suggests that the a priori design we naturally practice becomes the subject of inquiry for the researcher to segment its

organization. Therefore, various methods were used to study parts of this organization or the organization as a whole, depending on their unique perspectives to make their study operational. Moreover, new perspectives have also been suggested to examine the interaction as a dynamic process in its entirety. Our methodology is influenced by the framework of De Jaegher and Di Paolo (2007), inclusive of both individual and social dynamics of the interaction as the subject of the study. Authors recommend that interaction dynamics be studied in their entirety without focusing exclusively on individual contributions or social elements. It is important to view the interaction together with the participants and include the process of social interaction itself. Early vocal interaction between dyads is evaluated through the lens of both observed and coordinated social behaviors. As Condon proposes, what is claimed to be natural is “neither discrete nor continuous” but is ordered enough to be studied (1976: 285).

### **3.2. Primary Intersubjectivity and Interactional Synchrony**

The character of the mother-infant relationship and its interactional qualities are also essential aspects to consider. One of the first research on these qualities was conducted by Tronick et al. in 1978, with the famous “still face” experiment. Researchers examined the reactions of the infants toward the attitude of their mothers. Infants expect appropriate answers from their mothers, and when the mothers keep their faces still, infants show disturbed manifestations like crying. Another critical and quite similar experiment was conducted by Murray and Trevarthen (1985) on the importance of synchronizing reactions and attunement. In this study, mothers were recorded previously, and the video was shown to infants later. The research demonstrates the discrepancy between the mother’s and baby’s reactions, and this interaction creates discontentment because the mother’s reactions are neither adequate nor synchronous to the infant’s motives. This experiment shows how important it is for the infant to receive well-timed and tuned signals from the caregiver.

Developmentally speaking, primary intersubjectivity is another vital function going hand in hand with communicative musicality. Trevarthen introduced the notion of “Primary Intersubjectivity” (1974, 1977, 1998), describing the infant’s capacity to

express and communicate with the mother during the first six months of life. This notion resulted from subtle observations of infant-mother couples in a micro-framed analysis. These observations lead him to uncover the motives of the infants to communicate with their mother, their ways of expressing themselves, how they react to their mother's behaviors, and how the dyad share this experience together (Trevarthen, 1977, 1993). This experience is highly synchronous in terms of rhythmical mutuality, involving anticipation and coordination between two communication partners, leading to emotional and interpersonal exchanges (Trevarthen, 1989, 1993). These qualities carry out an early musicality among infants. The musicality of voice and body already communicates intention even before owning proper linguistic meaningfulness (Trevarthen & Malloch, 2002).

The nonverbal features of communication and interactive synchrony lead researchers to study the neurobiological roots of early social development through primary intersubjectivity. Trevarthen (1996, 2001) highlights the importance of right hemisphere functions' activation, i.e., as a response to the maternal prosody, and how these socio-emotional aspects affect intersubjective regulation. Much research into the physiological foundations supports the communicational settings of primary intersubjectivity. Different modalities of primary intersubjectivity, auditory, visual, and tactile, among others, lead two brains to communicate, while the infant's developing brain uncovers basic functions to be the foundations of future achievements. Infants' well-known early sensitivity to perceive prosodic aspects of maternal voice (e.g., De Casper & Fifer, 1980), as well as their early capacity even to imitate certain communicational signs (e.g., Meltzoff & Moore, 1977, 1992, 1994) already shows us infant captivity for reciprocity of early communication.

This reciprocity, often called synchrony, has been the subject of an investigation by many authors on aspects such as expressive manners (Brazelton et al., 1974; Jaffe et al., 2001; Stern, 1982), biological terms (Feldman, 2007; Porges, 2011) or matching between different modalities, e.g., leg movements and speech, (Condon & Sander, 1974).



The emotional and social aspects of early intersubjectivity turned attention to the question of emotional regulation and how the mother-infant dyad works on calibrating internal states. Porges (2011) highlights the vagus nerve's role in responding to environmental stress by regulating internal balance (homeostasis). He suggests that developmental activity of the vagus nerve, the vagal tone, during early infancy is the primary step in the building hierarchy of self-regulation. Feldman and colleagues (Feldman et al., 2010; Feldman, 2007) studied the impact of the emotional regulation process by monitoring biological rhythms by measuring levels of oxytocin or cortisol along with vagal tone. These studies demonstrate that infants' physiological responses to stress depend on the interactive synchrony formed with the mother. While the mother communicates on at least one level, the infant's biorhythm comes closer to a homeostatic level. This synchrony arises as a sense of rhythm through the vibrations perceived on many corporal levels. Stern (1985/2000) suggests that it is time for the infant to feel the basic temporal structure of verbal flow along with bodily movements, which is considered a baseline movement for the psychological basis of time experience and as a regulation proto-tool.

Dissanayake (2000) describes synchrony and turn-taking as units of modulation and regulation of the affective state. Therefore, synchrony is not only an agent of regulation but also serves to form a primary ground for infants to express their identity through vocalizations or movements, which creates a modulation during the interaction. For example, an infant's contribution modulates the interaction flow with the mother (Delafield-Butt & Trevarthen, 2015). This intervention causes a distinct energy increase in both parties. Stern (1995) suggested that narrative-like sequences of synchrony and regulation during early mother-infant interaction stem from biological needs (e.g., hunger) triggered by the infant and fade out by the response of the mother (e.g., breastfeeding). Susanne Langer says, "to feel is to do something" (1988:7); in other words, feelings afford expressive action. To feel is also to be moved by the other in all possible manners, and that movement contains rhythm and flow by nature. The aspects of primary intersubjectivity are, in fact, moving infants' observable expressions but also bear internal changes during early periods of life. Movements generated by the "intrinsic motive pulse" (Trevarthen, 2000) originate in early interaction as a

potential basis for “vitality affects” (Stern, 1985/2000), a dynamic process of forms of development and alterations throughout a lifetime (Stern, 2010).

### **3.3. Neurocognitive Roots of Musicality**

In addition to the adjustments in neurobiological levels, musicality also evokes perceptual signals linked to the activation on the brain level. Many authors have already studied an analogy between structures of music and language from a neuroscientific perspective (Koelsch, 2011; Patel, 2011, 2008; Peretz et al., 2015), showing that intersected areas of the brain are affected by music and language, including an adaptive plasticity hypothesis operated by music for the sake of processing speech (Patel, 2011). Researchers show that temporal regions of the brain react strongly to musical stimuli compared to speech and non-vocal stimuli (Peretz et al., 2015); the left auditory cortex responds more to linguistic phonemes, whereas the right auditory cortex is more responsive to musical and speech melodies (Koelsch, 2011).

In a series of research on neural populations affected by different sound excerpts, Norman-Haignere et al. (2015) found different neural pathways involving speech and music stimuli in the auditory cortex. Researchers have tested 165 natural sound samples, including speech, ringing, and breathing, to identify specific cortical response selectivity by categorizing different sound components, such as frequency and pitch. Results indicate different neural populations selective for speech and music in the auditory cortex. In a recently published article, the same research team also identified a particular neural population responding exclusively to singing (Norman-Haignere et al., 2022a) but not speech or instrumental music, confirming their previous findings. This study suggests that distinct neural populations represent different components of music within the auditory cortex. Researchers also found hierarchical organizations of sound in the auditory cortex over different timescales (from 50 to 400 ms); short scales represent acoustic components of sound, while long scales (over 200 ms) indicate category-selective neural populations. Results illustrate how human beings integrate musical stimuli using various timescales and how the brain organizes sound stimuli by time-framing to conceive meaning (Norman-Haignere et al., 2022b).

These studies show that different musical and sound stimuli have specific representative neural populations in the auditory cortex and that some neural populations are dedicated to both unique and unified spectral components of sound and music. In other words, research on neural populations shows that the brain functions as a pattern recognition device to capture various auditory stimuli. A cluster of neurons is an agent of abstract contingency.

Many cognitive studies have been dedicated to the structural organization of both music and language. Jackendoff and Lerdahl (2006) examined cognitive structures applying to music through elements of rhythmic and melodic organization, which are shared with language, along with the affect incited. Authors pointed out that both music and language have a structural organization grouping events (i.e., based on rhythmic and tonal components) depending on their tension-release flow (affective framing). Furthermore, recent research proceeded from the question of what stories we imagine during music listening in a comparative study with participants from different cities in the United States and rural China. This study on instrumental music listening shows evidence of narrative listening across cultures. Nevertheless, people growing up in the same culture use similar patterns to imagine these wordless musical stories. Similar story forms were pictured among U.S. citizens, using similar words for their descriptions. However, in cross-comparison, participants from China differed from the U.S. participants in terms of imagined stories, but their envisioned stories shared similarities within the culture. Results show that shared cultural background influences common patterns of narrative imagining while listening to instrumental music (Margulis et al., 2022).

In a similar study, researchers investigated event-structure responses to music by monitoring default mode network (DMN) activity (Williams et al., 2022), a network known as an operative for self-reflection and self-awareness during activities such as daydreaming and musical listening (Reybrouck, Vuust & Brattico, 2018). Previous work on event structuring showed evidence of the involvement of these regions during movie watching (Baldassano et al., 2018, 2017), leading researchers to test similar activation during the listening of 16 musical samples from classical and jazz music. Data was collected by both brain imaging methods and listeners' acoustic observations.

Results show a match of event segmentation by comparing fMRI records of participants and observer connotations based on temporal limits of segments (Williams et al., 2022). Participants spotted similar boundary markers between segments (events) across shifts in the musical flow. Results also confirm that the same brain areas are activated in response to music-listening and movie-watching event segmentation (Williams et al., 2022; Baldassano et al., 2018).

Growing evidence following recent brain imaging studies increasingly supports the evidence from previous observational studies on how processes of narrativization occur in the human mind.

### **3.4. Musical Narratives: Saying something musically...**

This part will look over the reflections and research on music and narratives from a larger perspective on music research and mother-infant interaction, attempting to define “musical narratives.”

Inspired by Chomsky, Lerdhal, and Jackendoff (1983) proposed a model depicting tonal music as a similar system to language by inquiring into musical syntax. These syntactic segments serve to analyze how musical units form an extensive structure and how these structures become explorable through the moments of tension and relaxation perceived by the audience across the segmentation of musical flow. Further research supported psychological representations, allowing us to recognize and perceive music in terms of metrical hierarchy (Palmer & Krumhansl, 1990) and segmentation of musical units (Clarke & Krumhansl, 1990).

Research on classical music interrogated the question of narrative structures based on joint accounts such as temporal organization defined through literary work. Imberty (2005, 2008) proposes that narratives in music define human time and represent a division from the time we live. His work essentially consists of the narrative dimension of music. Repetition and variation appear as recurrent patterns in music, and they lay out the structuring of the segmentation of the time of a musical work. Maus, in his work entitled “Music as Narrative,” expresses his thoughts on similarities between classical music and literary narrative (1991), asking the question, “Does the

music encourage this comparison?” is the hierarchical structure of sequences enough to describe music as narrative? He tries to answer this question by asking how listeners perceive music and describing the “musical plot.” The answer comes from Newcomb’s work reporting that listeners think of “thematic units partly as characters in a narrative” (Newcomb, 1987, cited in Maus, 1991) or as actions (Cone, 1974, cited in Maus, 1991). According to this assessment, if we consider music as action, the idea of the narrative comes naturally since actions already have a starting point and an ending. Nattiez (2013) further discussed music as a proto-narrative, inspired by the work of Daniel Stern (1985/2000) and emphasizing the symbolic distinction between music and language, affirming that music never tells a story the way that linguistic expression can do.

The distinction between music composition and improvisation requires other perspectives to question the narrative process and organization. Can we call improvisation a musical discourse and musical composition a musical story?

Daniel Stern allegorizes mother-infant interaction to a jazz duet. He illustrates the variation following regular patterns as expressive, like in jazz music; variations breaking the expectation create excitement (Stern, 1982). The musicality of the mother-infant interaction opens the way to new research comparing it to jazz improvisation (Schögler, 1999, 2002) and defining similar patterns of interaction as observed in communication between musicians (Miell, MacDonald & Hargreaves, 2005).

Musical improvisation is basically an *improvisus*<sup>1</sup> performance created by two or more musicians and based on listening to each other, finding a common beat, and weaving around musical motifs by adjusting each other. Each musician contributes differently, creating change in the flow, nonetheless maintaining a baseline texture.

Keith Sawyer (1992) described improvisation as a real-time interactive process and reported that jazz musicians describe it as a conversation. Berliner (1994) also documented that musicians illustrate the process of improvisation as “conversation-likely,” and Monson (1996) stated further that they picture improvisation as talking or

---

<sup>1</sup> Lat. unexpected

“saying something.” The word “playing” is often replaced by “talking” or “saying.” Theorizing musical narratives require taking into account their interactive and improvisational basis. Ole Kühl reveals that jazz improvisation is organized in a figure/ground relation<sup>2</sup>; musical phrase as figure and harmonic changes as ground, and they are influenced socially by the place in which they play and with whom they play (2007).

Gratier (2008) worked on the interaction between two jazz musicians while improvising. Musicians were recorded during improvisation sessions, described as sharing a common ground through which they delve into individual and shared musical backgrounds with highly skilled performances. Some interesting patterns have been reported in this work to explore how the communication and negotiation processes evolve between these two musicians to create a work of art. These patterns are discussed by putting the notion of “grounding” (Clark & Brennan, 1991) into the center, sharing mutual knowledge through coordination of content and process:

Repetition and variation appear as fundamental elements of musical improvisation. Repetition contains forms of imitation, a common pattern both musicians use, and a standard to uphold common ground. Variation introduces new forms stemming from regular and repetitive forms, creates a progressive shift in the musical flow, and brings creative inspiration through framed exploration. The repetition of musical motifs generates other patterns, like *mirroring* when one musician follows the pattern that the other musician proposes. In the case of *rhythmic matching*, musicians are on the same tightrope by simultaneously playing accented notes and beats. *Punctuation* appears as a form of interjections of intention and anticipation from one musician to the other. These motifs require attention from the musician, who will predict and later create variations upon them. Completing each other’s motifs, *completion* or playing the exact same notes, and *synchronization* are also fundamental elements of improvisation and evidence of grounding. These elements, just like in verbal communication, show us that these two musicians

---

<sup>2</sup> The author refers here to the work of Ungerer & Schmid (1996, cited in Kühl, 2003) on figure/ground relations.

understand each other nonverbally and express themselves in tune. Knowing each other's musical styles certainly helps them to maintain and progress this interaction.

To continue the example of jazz music, the groove might also be seen as a grounding device and a narrative pattern. A groove moving in a repetitive rhythmic arrangement can also give a climactic effect. Rhythmic narratives have regulation qualities and contain "arousal qualities" of the groove (Pressing 2002, cited in Kühl, 2007).

Another form of interaction is observed between the composer becoming the narrator and the listeners as the narrator's audience, which we also observe in literary and oral narrative cases. Literary narratives are already accomplished like composed music works but arouse different emotions among readers and listeners. However, sometimes listeners become involved in the creation process of a "written" musical work, as we see in the example of 4' 33" by John Cage. Cage provided some instructions to the musicians on stage, composed of three identical units named "Tacet." Following the instructions, musicians do not contribute by playing instruments, but the audience and the environment do. The audience then becomes involved as much as the musicians. Consequently, each performance is unique and constantly rewritten or, in this case, re-performed. The authenticity of 4' 33" redefines silence in music and affords musical freedom to the creative process.

Referring to improvisation sessions of the John Coltrane Quartet, Iyer (2004) described the dynamic process of creation during these sessions: "The story is revealed not as a simple linear narrative, but as a fractured, exploded one [...] this process of examining the puzzling shards of these exploded narratives that we may reveal a mosaic with a discernible underlying pattern." (2004: 395). Iyer's approach here considers that different modes of narrative-making are possible, especially in the case of improvisation.

Duranti and Burrell (2004) reported that jazz musicians are perpetually discovering new harmonic or melodic ways to express themselves while improvising. This discovery contemplates the realization of the musical self through emotional expressions. These expressions communicate cultural signatures linked to their

musically traditional background and community. Hearing what the others play as a collective, listening to the entire creation, and rethinking it are the processes realized in real-time and have an effect of constant self-inquiry. Discovery of the self through sound and music is familiar to us from the young infant's world. Cultural background reflected in musical expression refers here, like in the example of "intermusicality"<sup>3</sup> of Monson (1996: 97), to the temporal and social dimensions during the improvisation. This collective creation in the present time mirrors individual pasts, ways of being and thinking, and social affiliations.

Similar patterns of self-expression and cultural signatures are also observed in early interaction. In 2004, Cowley, Moodley, and Fiori-Cowley published an article on how culture-specific signs that caregiver uses are mirrored in the infant's behavior at three months. The interaction process involving joint activity opens ways for infants to access "meaning" on the one hand. On the other hand, they are also exposed to culture-specific forms of expression. Evidence shows that infants not only express but also grasp these cultural motifs. By this, the process of semiogenesis begins at a very early age.

### **3.5. Pursuit of characterizing musical narratives**

In previous chapters, we presented different narrative forms with a relatively similar structure representing an introduction, a development part, and a conclusion. Since Aristotle, artistic expressions have inspired and driven attempts at narrative portrayal. First, the literary arts were the primary focus of the investigation into the narrative structure. Then, linguistic and anthropological studies allowed us to contemplate narratives as everyday life activities. In psychology, reflections on narrative activity enact how we sketch ourselves and our interactions with others. Consideration of diverse artistic forms of expression as the narrative is involved in many theorists' work, such as Dissanayake (2000), Ochs & Capps (1996), and Chatman (1978).

Research on oral narratives opens the way for contemplating the musicality of expression as a narrative. Therefore, it is more challenging to illustrate the structure of

---

<sup>3</sup> Monson describes intermusicality as "something like intertextuality in sound." (1996, p.97).



musical narratives; it needs a more flexible structural schema but still follows the basic narrative structure (introduction-development-conclusion). However, diverse sequential streams make it difficult to anticipate the narrative organization as we observe in the case of atonal music or sometimes in the making of improvisation.

The narrative structure is known for its linearity, going toward a direction. As it progresses through time, the fellowship of structural units makes the narrative an entity. The alliance between these units generates meaning through connection. However, some components of narrative, especially those we encounter in forms of expression outside of literary realms, are also described, emphasizing the existence of early occurrences and essential units. These narrative forms and units are variously called “proto-narrative envelopes” (Stern, 1992, 2000), which are basically time envelopes (Stern, 1992), “pre-narrative structure of experience,” or “narrative form of time” (Ricoeur, 1983), “lines of dramatic tension” (Labov, 1972), and “vitality contours” (Stern, 1999). Those diverse descriptions commonly reflect the idea that narrative thinking is an essential universal human capacity (Stern, 1998; Bruner, 1990).

To rebound from our initial question about whether a narrative is a structure or process, we will attempt to review this question through the lens of different musical genres and from the perspective of the creator and the audience. Since Propp, narrative researchers have worked on structural analysis, describing narrative as a model, trying to figure out how it takes shape, and noting recurrent patterns. Therefore, music theorists working on Western classical music define it as “narrative” (Imberty, 2005). Several authors have also attempted to theorize music, like Almén (2008) as narrative or Levinson as dramatic (2004), prior to Schoenberg’s atonal freedom and musical imagination making an appearance. Yet, in jazz or improvisational settings, narrative forms in music do not appear as a structure but as a process or a performance. It appears to be a pattern in progress. Following this model, we can consider that “the basic idea” of narrative (Stern, 1999) grows through phrases or chunks. We can rather call them *floating narratives*. Despite the disorganization of the information we are exposed to, we, the people, try to narrativize everything; human cognition is urged to organize any external input to make sense of it (Bruner, 1990).

This need also appears in music through expectations of harmony and recognition of a theme (Meyer, 1956).

In the previous chapter, we reviewed the first musical encounters of human infants at the very beginning of life. First, being exposed to sound, mainly the maternal voice and tongue, sketches out the basis of sound perception for the infant. Second, shortly after birth, infants start to produce sounds and vocalizations that carry out the traces of that initial sound exposure. The rapid development of infants, especially the increasing capacity to communicate, changes their role during interactions. First, an infant is just the audience to sound and musicality, then rapidly becomes the participant as their vocal quality grows and, finally, becomes engaged in the co-creation of interaction with the primary partner. From the perspective of communicative musicality, an infant gradually turns into a musical agent, both a listener and creator. This process, by its similarity, shows us how researchers studying mother-infant interaction are also interested in analyzing the interaction between jazz musicians during improvisation.

Atonal music, however, offers a different point of view to music creators and challenges our need to organize as listeners. Schoenberg's work shows how musical expression can differ, with infinite possibilities of thinking and creating, and how the structural patterns can be used dispersedly as puzzle pieces in the whole musical work. Atonal music, also described as dissonant, represents a deviation from the traditional harmonic structure of tonal music, which carries out the elements of music's narrative account. Samson (1974) suggested that "an analytical approach to 'atonality' should begin by accepting a wide range of interacting functions in an *oeuvre* where tradition and innovation are inextricably interwoven, and the emphasis lies as much on the *exploration* of new resources as on their organization." (1974: 16). Instead, he proposed to look at the similarities, such as phrasing and musical forms, while analyzing atonal music rather than focusing on apparent dissimilarities through the lens of tonal music. Simon also reported that Schoenberg also had a "need to organize atonality" to create new forms and meanings, despite the changes of centrality progressing through tones in his compositions. He emphasized that Schoenberg's "continuous variation" also included some repetition to validate his search for a

thematic ground (1974: 19). Phrasing and forming using atonal principles pattern the so-called musical narrative differently, despite the absence of salient narrative components found in tonal music.

Less study is dedicated to infant and child responses to atonal music excerpts, mostly suggesting a tendency for tonality. Van Puyvelde et al. (2014) measured cardiorespiratory effects on tonal and atonal music pieces of mothers and 3-month-old infants. Results showed infants' preference toward tonal stimulus by indicating increased vagal activity signifying a more regulatory process than atonal stimulus. A previous study on six-month-old infants measuring visual fixation coupled with auditory exposure showed similar results for their preferences for consonant intervals compared to dissonant intervals (Trainor & Heinmiller, 1998). A similar experimental design has been used, indicating the preference for consonance over dissonance in two-day-old newborns of both deaf and hearing parents (Masataka, 2006). Later, Masataka (2017) suggested that atonal music listening is more demanding in aesthetic terms in a study comparing neurotypical and autistic children (4 to 7 years old). This paper displays the preference of autistic children for atonal music (Schoenberg and Albinoni) compared to their neurotypical peers and discusses autistic children's increased capacity for music, such as their remarkable music memory and absolute pitch mastery, to illustrate this tendency. Further research is needed to describe the reasons for this preference.

In this chapter, we tried to show analogies between musical improvisation and conversation through creation, spontaneity, and culturally grounded signs. In the context of mother-infant interaction, we suggest that the infant creates an early sense of identity and communication style through *narrative socialization* throughout the interactional moments. Musicality is a way to convey emotions and intentions and draw up a musical culture in an infant's world (Flohr & Trevarthen, 2007).

Human creativity is not limited to artistic expression, and we constantly find ourselves needing to express our feelings or thoughts in many interactive ways. We observe different portrayals of narratives in musical expression by looking at different musical genres. The organization of expression can be described in two phases:

progressing through processes and becoming a structure. It is also essential to consider the distinction between our ways of creating or expressing ourselves and our need to organize to understand this expression. From this perspective, early interactional settings seem to be an initial process for the infant to enter the narrative realm. This speaks for an abstract primary representation of a future elaborate organization. Over a reasonable period of time, perhaps until the complete acquisition of language, human infants go through narrative processes to expand their ways of expression and understanding.

## Chapter 4: Meaning: At the Crossroads of Words and Melodies

That was when I learned that words are no good;  
that words don't ever fit even what they are trying to say at.

William Faulkner, *As I Lay Dying*, 1930.

This chapter is dedicated to contemplating meaning in musicality as well as musical meaning. We review theoretical approaches from multiple disciplines, such as linguistics and semiotics, among others, to conceive how meaning appears in the early stages of life as proto-forms and how it differs from linguistic meaning. . It is crucial to review structured and improvised manners of conveying meaning, although this raises more questions, while maintaining focus on music and language are the two core meaning-making processes.

The pioneer of Structuralism, Saussure (1916), introduced a distinction between *parole* (effective realization, actual speech) and *langue* (system, language structure). His work does not focus on the use of language but rather on its underlying system by emphasizing the sign as a basic unit of the language system. To present his binary model, Saussure described the sign function; through the arbitrary link of a signified (the concept) and a signifier (the sound pattern). This link gives the sign its signification (not the sense or meaning) only within the respective language system, not by its reference but by being differential from the other signs of the system. However, Benveniste (1939) described this link as necessary by invoking a referent, an object, which is an absent dimension in Saussure's linguistic system. Sapir (1929) explained phonetic symbolism in language by binary distinction. He described associative symbolism as an utterly arbitrary forming of phonetic components (similar to Saussure) and expressive symbolism, which relies on phonetic differences in actual speech. Roland Barthes (1953) also presented a model including *écriture* (translated as writing), style, and language, and privileged writing as something "beyond language." Barthes

(1953) was particularly interested in writing itself, in its structure rather than the meaning.

In “Six Lectures on Sound and Meaning” (1942), Jakobson emphasized the relationship between sound and the meaning of words through the example of Edgar Allan Poe’s poem “The Raven.” Jakobson pointed out the repetition of the word “Nevermore” in the poem and its general and contextual meanings. Along with Poe, he indicates the acoustic quality of the repetition of this word, which is, for Poe, the core of the poem’s inspiration. Between Benveniste’s necessary relation between signifier and signified (1939) and de Saussure’s arbitrary relationship (1916), and Sapir’s sound symbolism (1929), Jakobson opens a new way of thinking about the acoustic pattern of a word through poetry.

Mead (1934) portrayed vocal gestures as a stimulus generating a response. He described the imitation process through the echo and the influence of the vocal gesture on the correspondent who is expected to give a response to that vocal gesture. When the second party imitates the vocal gesture, it is considered significant. By this, Mead distinguished “conversation of gestures” produced by birds or babies for themselves from “conversation of significant gestures” as linguistic interaction. A vocal gesture makes sense when it signifies the same meaning for both parties and when it incites a response from the addressee.

Whereas in interactional settings, there is a constant process of meaning-making through creating a “common ground” that is dynamic and affected by mutual beliefs, assumptions, and knowledge that Clark & Brennan (1991) called “grounding.” This common ground contains a wide range of cultural know-how but also individual signatures blended in mutual expression. In other words, the grounding process is the adjustment of meanings, both conveyed and understood. These exchanges occur in the case of dialogues, in the shape of turn-taking with certain simultaneity in sequences following each other without interruption (Sacks, Schegloff, & Jefferson, 1974).

From a perspective including narrative forms, Bruner (1990) presented narrative interpretation as a meaning-making process through “negotiating and renegotiating meanings” through recalling and reconstructing memory simultaneously.

This process certainly includes a cultural background besides the impact of the evolution of linguistic capacities. According to Tomasello (2003), extracting those kinds of patterns serve to organize the input of information of all kind, especially of abstract linguistic constructions. Bruner described the production of this process as “format,” in parallel terms with Kühl described organization units as “chunks” (2011), and Lakoff and Johnson described abstract structures of embodied experience (1999).

According to many studies on mother-infant interaction, we are gradually weaving our abilities of social bonding and communication, which starts with the intersubjective sharing of emotions between the infant and the caregiver. This exchange is multimodal – visual, tactile, vocal - unified in the infant’s mind, yet still abstractly. This process influenced Ole Kühl in conceiving musical sign and musical meaning, a two-step model describing both signification and meaning, emphasizing that sound embodies movement. For Kühl (2011), a musical gesture is at first a purely cognitive entity, a representation in mind, and becomes an internalized gesture later on (i.e., a mentally represented gesture). Kühl creates a model adapting the Saussurean sign function and his dyadic model. In the first step, he describes the musical sign function, the musical phrase functioning through a sound pattern, and labelling musical gesture as a concept. In the second step, the musical gesture is taken as a sign providing an embodied meaning across its functions; “implied movement” as signifier and “emotional state” as signified, referencing Saussure’s description of the sound pattern, not being a physical source but the psychological conception of the hearer (Saussure, 1916). The first model represents signification on a primal basis, and the second model embodies a sense of social belonging. Musical gesture indicates an essential and basic pattern, “an intentional and meaningful act” (Kühl, 2007: 145), portraying an abstract, expanded level, contrasting to words that are precisely salient. He also suggests that music is as pertinent as language in terms of meaning.

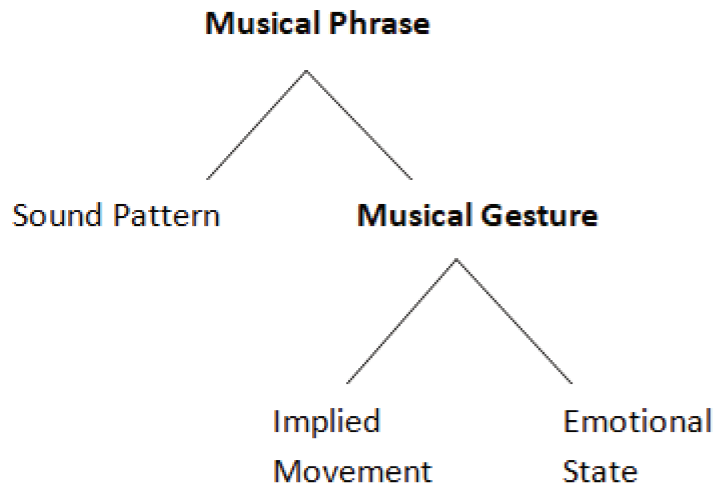


Figure 2. *The Gesture as Sign* (Kühl, 2011, p. 128)

While looking closer at jazz improvisations, Ole Kühl (2007) says that musicians interacting through improvisation already have the rhythmic and pitch contours' representations. Based on this representation, they focus on two levels: the first one is the details of phrasing, and the second one is the structure of the story they are going to tell.

Keith Sawyer (2002) worked on how collaborative improvisation of child plays takes the form of a narrative. In this context, the narrative-making process has an unpredictable nature developing step-by-step, is collaborated and integrated into the social context. Sawyer analyzed every level of improvised child play development and listed strategies children use during play. To shift the play's storyline, children sometimes talk out of character and change the strategy of the play, which he calls *metaplay*, or they talk in character, adopting a *dialogic strategy* (Bakhtin, 1981, cited in Sawyer, 2002) to negotiate or guide the storyline of the play. Sawyer says that in an improvisational enactment case, labeling the play as a narrative requires the transcription of the act to analyze it afterward, which contradicts the essence of improvisation.

In "Narratives from the Crib" (2006), Katherine Nelson tells the story of a girl's monologues before sleep. These monologues present us with how she sees and



describes her inner life and the world around her and also provide us with information about a two-year-old girl's social and linguistic realm. This work shows how a child creates a basic narrative framework by herself. Sawyer and Nelson's works are examples of how children are capable of meaning-making and creating stories using their linguistic and symbolic capacities.

Researchers suggest that language and music in the same communicative system have common patterns expressed phonologically and meaningfully (Brown, 2000). Speech conveys emotion through its melody and rhythm, the prosodic components of emotion in speech (Scherer, 1986, 2003). However, music gives us the possibility of multiple interpretations. Music could carry different meanings to the composer or performer, the expressive side, and to various listeners, the receptive side (Kühl, 2007). It might also generate a common meaning for numerous people. This elasticity of meaningfulness also depends on the context. Ian Cross calls this "floating intentionality" (1999, 2003), emphasizing the meaning coming from the context and the attribution of meaning to the context. This might be the crucial point separating music from language despite many similarities.

Meaning that music generates does not only have one literal meaning like words of a language. Somehow, it still gives us the possibility to anticipate what comes next in terms of rhythm or melody while listening to a musical composition (Narmour, 1990; Meyer, 1973, cited in Imberty, 2008, p.134). Some of these anticipations become attained, some of them not. Attained expectations might have common realizations that Meyer calls "evident meanings" (Meyer, 1956, p.37). He suggests this is possible when the listener decodes the relationships between musical motives. Gradually, the unified illustration of this chain of causality could become a referential point to be represented throughout the development of the musical piece.

In music, themes are represented. However, they are not reduced to a single topic word; they also represent connotations. This connotative complexity might bear a singular meaning or topic for each listener (Meyer, 1956). Words have explicit meanings; however, "musical meaning is *fluid*" (Kühl, 2007, p.37). Ole Kühl reports that there are objective elements in music that we can assume, but it also contains many

subjective elements that we can interpret individually differently (2007). These subjective interpretations can be reduced among the same cultural groups. These three poles are combined in a linear form of music's "semantic axis" (2007, p. 40): Cultural and objective features represent shared meaning, yet subjective features remain fluid. Köhl, presenting musical semantics and preverbal semantics in the same wavelength as proto-semantics, provides us with a significant description of how we should interpret meaning. He describes musical semantics as a non-linear form, as opposed to the linear form of language.

As Köhl pointed out, music owns an implicit movement. For an infant who does not walk yet, the perception of rhythm comes from the proprioception of his own heartbeat and breathing or the movement of different body parts. Köhl says that we represent "pulse and meter mentally as internalized walking" (2007, p.184), likewise what Brandt exquisitely described as "the private dancer" (2004), an act of mentally dancing while imagining music. The acts of walking or dancing can be seen as narrative activities, corresponding to Genette's description of minimal narratives (1972).

Intonational variations afford emotional significance in music and language (Köhl, 2007). Fonagy (1983) stated that the transformation of speech through intonation, the paralinguistic trend, is ample to transmit emotion and is captured even if the addressee does not speak the language. In the speech, musical components are articulated in syllabic and syntactic entities combined to become phrasing units. Brandt says, "[...] in ordinary speech, we spontaneously "sing" the overall structure of our grammatical sentences in accordance with the intersubjective circumstances and our purpose." (2009, p.32). This spontaneous singing appears more saliently in the context of mother-infant vocal interaction. In "The Poetics of Babytalk," Miall & Dissanayake (2003) showed the poetic texture of a mother's speech to her 8-week-old baby, and these features generate mutual emotional expressions by both partners. This early age also corresponds to the period in which infants' sensitivity to time patterns appears (Jaffe et al., 2001). These pattern sketches allow them to engage fully in a conversation, and mother and baby become "attuned" around 6-8 weeks old (Stern, 1985/2000) by shared coordination and expressiveness. Meaning is embedded in the musicality and attunement of two bodies.

Theoretically, the intensity of sound is what Daniel Stern described as “vitality affects” (1999). The change in intensity over time corresponds to our vitality. It also refers to the changes in internal body sensations like heartbeat and pulsations. Susanne Langer defined “feeling forms” (1953); as feeling in its extensive meaning, from physical sensations to inner tension. For Langer, it is quite inconceivable to describe feelings in words. Words can only draw a broader frame to describe feelings. Hence, words remain restricted from fully expressing the fluctuations of an inner feeling. Langer says, “[...] *But there is no language to describe just how one joy differs, sometimes radically, from another. The real nature of feeling is something language as such – as discursive symbolism – cannot render.*” (1966, p.9, italics added). As a case in point, Milan Kundera strives to define a word of feeling, existing only in the Czech language, “*litost*,” in his book “The Book of Laughter and Forgetting,” through some examples of situations which attempts to generate the feeling that the word represents. Langer defined feeling as “a pattern of organically interdependent and interdetermined tensions and resolutions” (1966, p. 9). The experience of mastering those feelings is possible only through repetition and elaboration of these forms. The combination of these forms is related to the expression and creation of a work of art, especially music. Form in Langer’s vocabulary should be considered a dynamic form in movement, a “living” form. One of the critical aspects of Langer’s reflection is the analogy between forms of feeling and music. There is constant change and evolution in how feelings are expanded and music flows. Besides, the analogy is interwoven through the expression of a work of art. However, Langer (1966) specifies that music is a symbol, and its symbolism is presentational but not discursive like language to contrast these two in her reflection precisely. When defining musical meaning, Langer says the musical symbol is different from the linguistic symbol, which does not refer to an exact definition of meaning. She does not reduce the expression through a piece of music to describe self-expression. Still, she defines it as the communication of feelings as a consequence of the composer's intention and will to communicate.

The idea of an implicit movement in musical meaning leads us to think of other concepts in the intersection. The concept of *affordances*, “*possible actions*,” was coined by Gibson (1977, 1979) in his Ecological Approach to Visual Perception theory.

Briefly, this concept represents the link between the environment and the organism and the potential motion that this interaction bears. Several authors asked what music affords, and the term “musical affordances” was introduced by some researchers (Clarke, 2005; Krueger, 2011a; Reybrouck, 2005) from different perspectives and interpretations. However, the term does not clearly point out what it means. Clarke says, “Music affords dancing, worship, co-ordinated working, persuasion, emotional catharsis, marching, foot-tapping, and myriad other activities of a perfectly tangible kind.” (2005: 38). In his book, he applied the concept of affordances to reconsider musical meaning through it. He detailed some examples of musical performances. This gives a high range of possibilities of meaning depending on context and listeners. The same multiplicity of meanings was previously discussed by Meyer (1956). Reybrouck (2005) focuses more on the relationship between the body and musical cognition. For him, musical experience and motor cognition are closely linked, so he emphasizes the pertinence of action in musical perception. Krueger (2011a) suggests that from the beginning of life, music affords a sonic world that also affords social coordination and emotional regulation.

Rethinking back to the definition given by Gibson (musical) affordances point out multiple links furnishing different levels of responses and meanings. In a young infant’s world, while discovering the environment, they also discover how to communicate with this environment. Exploring their vocal capacity by producing sound, proprioception of vocal production, and the feedback given to their vocalizations encourages them to repeat and vary their own voice (Imberty, 1981). According to the interaction between the infant and caregiver and its musicality, musical affordances could be pictured through synchronized movements of the infant’s vocal production and attuned co-vocalizations or facial expressions.

To conclude this topic, we epitomize this part by reconsidering the meaning in narrative forms. As a narrative form, music presents infinite possibilities for meaning-making and interpretation. This is, therefore, possible through variations and changes within the musical work. Langer (1942) describes the primary characteristic of music as articulating feelings, through which it may or may not generate a cathartic effect. As

frequently described, the meaningful effect comes from music's implicit components arousing emotions and molding movement embedded in the style and expression.

The infant's contribution amplifies the narrative effect and meaning in the mother-infant interaction context regarding the musicality of early interaction. Moments of co-constructed units of melodic contours, intensity, or rapid movements of changing voices intensify the "feeling of form." Thus, we consider this early stage of protosemiosis through musicality as a narrative practice of meaning-making.

## CHAPTER 5: METHODS

### **5.1. Rationale and aims of the study**

This present research brings, in terms of innovation, an approach to making an operational descriptive analysis based on observation of narrative patterns in mother-infant vocal interaction. Sound-based musical narratives in mother-infant vocal interaction have been described qualitatively in previous studies demonstrating detailed descriptive analysis over chosen examples. (Delafield-Butt & Trevarthen, 2015; Gratier & Trevarthen, 2008; Malloch, 1999; Malloch & Trevarthen, 2009). Before these studies, authors described this pattern theoretically as a *proto-narrative envelope* (Stern, 1985) or as *joint action formats* (Bruner, 1990). However, no large-scale quantitative analyses have been conducted until very recently.

Research on oral narratives helped us conceive a basic methodology for identifying such episodes or structures in everyday adult-infant social interaction. However, a novel approach needed to be devised as mother-infant interaction is not purely linguistic. Thus, the interaction's nature differs from oral narratives in adult to adult interaction. Our analysis is based on acoustic features of the interaction as well as linguistic components. We mainly focused on acoustic elements of both infant vocalizations and maternal speech. That is to say, over this particular interactional style, we attempt to define a minimal set of acoustic and linguistic parameters for identifying narrative episodes in mother-infant vocal interaction occurring during relaxed face-to-face encounters. We base our analysis on the temporal organization of vocal units taken as conversational micro-units interactively produced by mothers and infants. The theory of Communicative Musicality (Malloch, 1999; Malloch & Trevarthen, 2009) provides a valuable theoretical grounding for setting up our empirical model. Nevertheless, we needed an exploratory phase of data analysis before outlining a stabilized methodological framework for quantitative analysis of vocal narratives of mother-infant interaction.

We chose to focus our analysis on vocal descriptors, including acoustic parameters and verbal content, even though narratives are not reduced to vocal expression but

involve corporal and facial expressions often synchronized with vocalizations (Delafield-Butt & Trevarthen, 2015). We also chose to focus on a restricted age, 2 to 3-month-olds, because the second month represents many ontogenetic developmental changes in infants' vocal repertoire (Gratier & Devouche, 2011). The second month is also a period when parents start to discern and describe their infants' vocalizations as "speech-like."

The first aim of this research is to identify patterns in spontaneous mother-infant vocal interaction that are analogous to other kinds of narrative organization in text, oral discourse, or music, based on previous qualitative studies.

The second aim of this research is to finely describe the acoustic parameters and verbal content categories of the identified samples obtained from two distinct corpuses (EVOC and VERONA, described in the Methods section). We have selected two syntactically and grammatically similar languages, French and Italian, belonging to the same linguistic rhythmic group of Latin origin from Europe, to detect similarities and distinguish specific narrative styles between the two cultures.

The third aim of this research is to operationalize a minimal set of acoustic parameters that serve to differentiate narrative from non-narrative episodes in spontaneous, everyday, relaxed vocal interaction.

## ***5.2. Research Questions***

Our main research question is whether very young infants coordinate their vocalizations not only within a turn-taking format, as has been shown in previous studies (Gratier et al., 2015; Rohlfing et al., 2020), but also within a narrative format. In other words, we wanted to know whether infants participate in constructing narrative episodes. Specific acoustic characteristics of vocal narrative speech would thus not only be identifiable by young infants but also afford their vocal participation.

Based on previous research on a variety of forms of oral narrative, including largely qualitative studies of infants' interaction, we anticipate that **narrative patterns** are composed of an initiating sound sequence establishing mutual engagement, followed by a sequence of growth where emotional engagement is heightened, forming a "line of dramatic tension" marked by elevated vocal pitch and/or intensity as well as simultaneous

or rapidly alternating vocalizations, and finally, by a decline indexed by a slowing pace and a drop in excitement. The studies we conduct seek to explore whether these patterns are present and recurrent in mother-infant vocal interaction.

Our central hypothesis is that narrative episodes are minimally defined by a significant rise and drop in fundamental frequency and intensity describing an arc-shaped dynamic form. Our second hypothesis based on the literature of narrative and social interaction is that infants will participate more and produce more speech-like vocalizations within a narrative episode, set up by the mother through her own narrative speech style, than within an episode that is not narratively set up. We think the narrative is a powerful motivational context for infants' early vocal development.



### **5.3. Sample selection from the EVOC corpus**

The EVOC study (Emergence of infant vocalization from 1<sup>st</sup> to 5<sup>th</sup> month, funded by the Fondation de France) was conducted at the LECD (Laboratoire Ethologie, Cognition, Développement) of Université Paris Nanterre by M. Gratier, E. Devouche and G. Apter. The corpus is composed of everyday home recordings of 24 infants and their families over a period of 4 to 5 months. Audio data were collected by parents who were given a high-fidelity audio recorder (Micro Track II by M-Audio) for use over the duration of the study, as well as instructions on recording procedures. All infants were born full-term and healthy. Both parents signed informed consent forms, and ethical approval was obtained for the study from the university ethics committee (CCP n°1450089). All mothers were native French speakers from middle or upper-middle-class socioeconomic backgrounds living in Paris or the suburbs of Paris. All mothers scored low on The Edinburgh Postnatal Depression Scale (Cox, Holden & Sagovsky, 1987).

The present research was based on careful listening, through spectrographic visualization using Praat version 5.4.02 (Boersma & Weenink, 2014) and Audacity 2.0.6<sup>4</sup>, of all the recordings made when infants were aged between 8 to 12 weeks. Around 100 audio sequences ranging from 3 to 10 minutes were analyzed (approximately 240 minutes total) for 16 of the 24 families having taken part in the EVOC corpus. In total, the collected sample is based on recordings from 8 boys and 8 girls (Mean age= 70.35 days, SD= 9.76). We specifically selected recordings where parents and infants were engaged in ongoing relaxed vocal interaction. We excluded recordings with low-quality sounds, infant crying, parent singing, or non-speech sounds in order to obtain a homogeneous sample of everyday social protoconversation.

### **5.4 Sample selection from the VERONA corpus**

A second corpus consists of Italian mother-infant dyads. Data collection was studied at the University of Verona under the supervision of Prof. Manuella Lavelli during my two-month residency as a visiting Ph.D. student. Twenty infants, all at three-month-

---

<sup>4</sup> Free download in <http://audacity.sourceforge.net/download/>

old, from Verona, Italy, and their mothers match our French corpus regarding socioeconomic background and education.

For the comparative study, Italian dyads were selected as a matched sample of native Italian-speaking mothers. All Italian infants were 3 months old precisely at the time of the recording. All recordings were made in the presence of an observer. We have analyzed a total of 200 minutes of audio recordings, each dyad having a precise ten-minute free interaction. Four dyads were discarded from analysis due to poor quality of audio recordings (e.g., recordings including environmental background noises). Only 16 Italian mother-infant dyads (8 boys, 8 girls) were selected for acoustic analysis to maintain optimum acoustic quality.

#### ***5.5. Criteria for identifying vocal interaction sequences in the EVOC and VERONA samples***

Audio recordings from the two corpora were listened to several times to detect features that determined whether mother and infant were engaged in vocal interaction.

- A sequence of vocal interaction begins after a 3000 ms pause in maternal speech.
- It has to include at least one infant vocalization
- It ends when neither mother nor infant utters any sound for at least 3000 ms.

We selected two samples of vocal interaction episodes defined by protoconversation, sound quality, and age of infants. Sequences were then segmented using Audacity 2.0.6 and relabelled as episodes of vocal interaction. From these samples, we identified narrative episodes based on the below chart. 64 narrative episodes (27.52 minutes in total) were selected with a mean duration of (25.80 sec  $\pm$  9.97) among French mother-infant dyads from the EVOC corpus. 60 non-narrative episodes (23.74 minutes in total duration) with a mean duration of 23.74 sec  $\pm$  9.87, matching in terms of duration to the correspondent EVOC narrative episode, were selected for the control group. 64 narrative episodes (23.18 minutes in total) (21.73 sec  $\pm$  8.58) were selected among the VERONA corpus as a sample of Italian mother-infant dyads. The quality of audio recordings

in a naturalistic environment is equivalent for both samples and intact for detailed acoustic analysis.

**5.6. Indicators for creating collections of narrative episodes from the EVOG and VERONA samples**

We devised a chart of indicators of narrative organization based on our analysis of the literature on oral social narrative forms. The chart presented in Table 1. is divided into four sections: three sequences identified phases of initiation, growth, and decline, and the fourth identified the overall global dynamic form. Each section includes indicators to pinpoint markers of narrative organization within the interaction sequence. The narrative indicator chart is later used as a coding diagram for further statistical analysis.

**Table 1**

*Indicator chart for narrative episodes*

INDICATOR CHART OF NARRATIVES		
<b>GLOBAL</b>	Topic Repetition	Bell-shaped Frequency and Intensity Curves
<b>DYNAMIC</b>	Internal Consistency	Repeated Melodic Contours
	Repeated Sound	Repeated Words
	Repeated Group of Words	
<b>INITIATION</b>	Ascending Contour	
	Short Pause Before the Beginning	
	Topic Change Compared to Previous Episode	
	Low Intensity	
<b>GROWTH</b>	Climax	
	Fast Vocal Turns	
	Co-Vocalization	
	Acceleration of Output	
<b>DECLINE</b>	Descending Contour	
	Lengthening of Syllables	
	Frequency and Intensity Fall	
	Decrease of Output	
	Long Pause at the End	

### **5.7. Criteria for identifying non-narrative episodes (EVOC sample only)**

Our study includes further analysis of a control group within French dyads for the purpose of having intra-subject control modalities. Some audio samples were selected within the same audio recordings, based on the same criteria for selecting sequences, with at least one infant vocalization to be observed. Sixty audio samples matching the duration of the correspondent narrative episode were chosen as control episodes.

Control episodes match the correspondent narrative episode in terms of duration and number of occurrences, except for four episodes discarded from the data for this comparative study for not having a similar control criterion. Thus, each narrative episode in the EVOC corpus has its control episode regarding matched pairs. We expect to see that these control episodes have longer pause durations between speakers. We also expect that control episodes have less infant vocal contribution and more maternal vocalizations in terms of number and duration within the same durational frame.

### **5.8. Collections of narrative and non-narrative episodes**

A total of 188 episodes were identified from two corpora. The French sample consists of 64 narrative and 60 control episodes, matched in terms of duration. 64 Italian narrative episodes were also identified for the narrative comparison study of both samples.

**Table 2**

<b>Corpus</b>	<b>Nationality</b>	<b>Narrative</b>	<b>No Dyads</b>	<b>No Episodes</b>
EVOC	French	Yes	16	64
EVOC	French	No	16	60
VERONA	Italian	Yes	16	64

## **5.9. Acoustic analysis of the collections**

Sound Analysis Pro (Tchernichovski et al., 2000), a computer-based spectral analysis software was used for semi-automatic coding of acoustic measures. This software which was initially created for the analysis of birdsong was found effective in studies of adult and infant vocalizations (Gratier et al., 2015; Lipkind et al., 2013). The procedure involved precise labelling of vocal units produced by mothers and infants. Once units are identified, duration, pitch, and intensity values are automatically tabulated.

### **5.9.1. Coding of vocal units**

(1) Maternal vocal units are defined as utterances where the mother's speech is continuous, allowing for short pauses of less than 300 ms within utterances. The verbal content of each utterance was transcribed.

(2) Infant vocal units are identified based on the continuous voiced sound produced by the infants, allowing for short pauses of less than 300 ms within them.

(3) Pauses superior to 300 ms and lesser than 3000 ms were coded

(4) Overlapping vocalizations are a superposition of simultaneous mother and infant vocalizations. Maternal utterances and infant vocalizations within overlapped units were coded differently from isolated vocalizations. Three types of overlapped vocalizations were coded:

- a) Mother and infant started and ended their vocalizations at the same time
- b) Mother initiated and infant joined in
- c) Infant initiated and mother joined in

(5) Sounds produced by infants, such as vegetative sounds like hiccups, breath, and sounds produced by mothers like coughs, were excluded from the coding procedure.

### **5.9.2. Acoustic Measures**

(1) Fundamental frequency ( $f_0$ ), measured in Hertz (Hz), is defined as the lowest frequency produced in a periodic swing.

(2) Intensity, the amplitude of a sound wave measured in decibels (dB), is defined as the chunk of energy transmitted per second over a unit.

(3) Duration of each vocal unit, as well as pauses

(4) Rate of production for each vocal unit and pauses

### ***5.9.3. Infant vocalization***

An additional measure of infant volubility (Iyer et al., 2016) was obtained. The proportion of infant vocalizations during vocal interaction is studied to observe whether some organizational patterns of interaction potentially impact the quantity of infant vocalizations. Measures of infant volubility of narrative and control episodes were compared based on infant vocal rate, percentage of time infants spent vocalizing, and the number of infant vocalizations. For comparison analysis, overlapped vocalizations were grouped together, discarding the onset differences.

## ***5. 10. Narrative analysis of verbal content of maternal utterances***

Narrative analysis was applied only to French data (EVOC sample) analyzing the verbal content of narrative and non-narrative episodes. The verbal content of maternal utterances was analyzed, including narrative qualities, based on four dimensions:

**Context** criterion includes references to place and time. Additionally, we analyzed tense use according to verbs in the mother's speech. Tense is coded differently from the rest of the data, based on the present, past, and future tenses and no tense use.

**Chronology** provides information on the temporal sequences of a narrative. In chronology, we looked for temporal order words (e.g., avant-après, bientôt, ce matin), repetition, elaboration, and end markers to pinpoint how these components appear along the timeline inside an episode.

**Theme** refers to the apparent topics of a narrative. Our analysis of this section is based on recurrent words, adjectives, questions, and interpretations. Interpretation includes both reported speech and direct interpretation of an infant's verbal or nonverbal behaviors.

**Agency** category includes coding of pronouns “you, me, us”, appellatives (e.g., mon coeur, ma puce), words of emotional attribution (e.g., tu es heureux), and attribution of cognition (e.g., tu comprends maintenant).

**Table 3**

*Verbal Content Coding Schema*

CONTEXT	CHRONOLOGY	THEME	AGENCY
Tense	Temporal Order	Recurrent Words	You
Place	Repetition	Adjectives	Me
Time	Elaboration	Questions	Us
	Ending Markers	Interpretation	Appellatives
			Attribution of Emotion
			Attribution of Cognition

We used a binary (Yes/No) scoring, a two-way close-ended answer based on whether the criterion exists or not, except for the "tense" criterion. The topic was determined based on the existence of a global theme specified by a word or a group of words. Items connected to each category represent a variable, and their interaction also corresponds to a hypothesis. Items forming a category represent an overall conceptual assembly. The verbal content analysis gives us a solid basis to link musicality with discourse in the interest of interpreting the items not only for their semantic meaning but also for their analogies with music (e.g., repetition and elaboration are also units practiced by musicians in the process of music-making).

**5.11. Acoustic analysis based on fundamental frequency and intensity graphs**

To further test the main hypothesis of this study, we conducted a study on the acoustic graphs of both French and Italian narrative episodes and French control episodes. We expected to observe ascending fundamental frequency and intensity values, reaching a maximal point while rising and then declining the acoustic energy towards the end (as in Figure 1), which we have called a narrative acoustic arc model. To

examine this hypothesis, one graph for fundamental frequency and one for intensity were drawn using respective values for each episode identified. Each couple of acoustic graphs was assessed by a blind coder taking into consideration the narrative arc model based on three criteria:

- (1) Comparing the equivalence of the beginning and the ending, which is expected to form a baseline in terms of acoustic values;
- (2) Existence of a climactic point marked by high frequency or intensity;
- (3) These two criteria should be valid for both fundamental frequency and intensity.

Binary coding (Fit/Not Fit) was applied for the graph couples that corresponded to the criteria. If the fundamental frequency graph corresponded but the intensity graph did not correspond to the criteria; it was coded as "no fit."

### **5.12. Hypotheses**

1. Specific patterns of pitch and intensity variation will constitute core acoustic features of vocal proto-narratives, with gradual increase and decrease of values across sequences of vocal interaction. A maximum pitch and intensity point is expected between the beginning and end of the episode.

2. The fundamental frequency and intensity range are wider in narrative episodes than in control episodes.

3. French and Italian dyads' narrative episodes will be characterized by the same minimal set of acoustic features.

4. Infants will vocalize more often in narrative episodes compared to control episodes.

5. Infants' vocalizations will be longer in narrative episodes than in control episodes.

6. Mothers will use more tense, agency and temporal order are expected in narrative episodes compare to control episodes.



### **5.13. Statistics**

Statistic analyses were done using SPSS version 20 and JASP version 0.13.1. An alpha level of .05 was used for all statistical tests.

Linear regression was conducted to predict the impact of the narrative effect on all types of vocalizations, including duration, vocal rate, and time spent vocalizing from French narrative and control episodes. Partial eta squared, as an index of the proportion of variance explained by the narrative outcome, is calculated to estimate the effect size (Cohen, 1988). Gender variables for infants are included in the model. Age is a controlled variable for French data. However, contingency analysis was conducted, including Cramer's V as an index of effect size, to see if any effect is revealed through 8-to-12-week intervals.

Logistic regression is conducted for binary coding data, including verbal content and narrative indicator chart, to calculate odds ratios and confidence intervals.

A principal component analysis using R was also applied to the indicators of the narrative chart.

## CHAPTER 6: RESULTS

### 6.1. Study 1: Narrative episodes in French and Italian samples

Several analyses have been conducted on the vocal interaction sequences identified in the EVOC (French) and VERONA (Italian) samples. The durations of the narrative episodes have been compared considering age and gender differences. Further, a comparative analysis of narrative indicators was operated along with principal component analysis (PCA) executed to determine primary narrative indicators based on the two datasets.

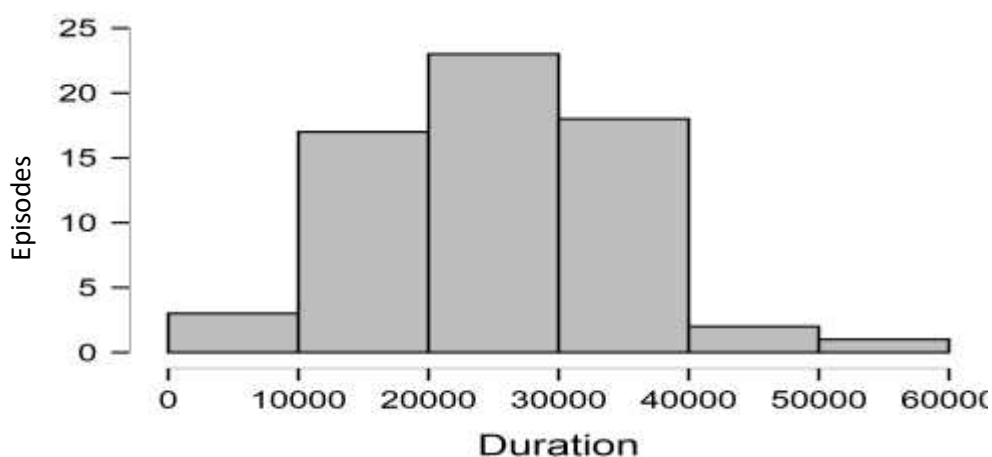
We expected to monitor small variations between maternal speech properties and infant vocal measures between two samples. We compared French and Italian narrative episodes to contextualize the differences in interactive styles through acoustic measures. Verbal content analysis has not been performed to cluster acoustic-only properties of the narrative organization.

#### 6.1.1. Duration of French and Italian narrative episodes

64 episodes matching the selection criteria (27.52 minutes in total) were selected with a mean duration of 25.80 sec  $\pm$  9.97, ranging in milliseconds from 7814.18 – 57629.21 among French mother-infant dyads for the collection of narrative episodes.

**Figure 3**

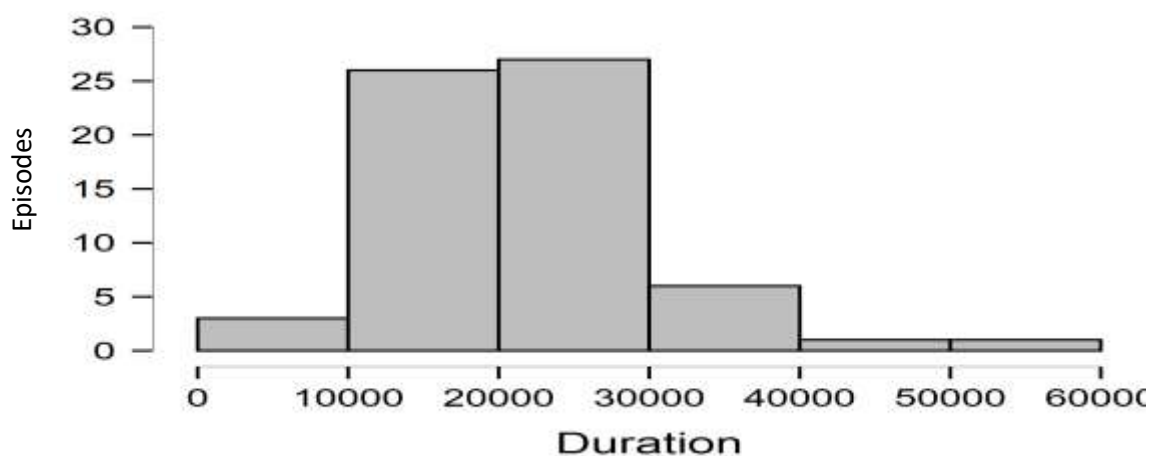
*Duration distribution (in milliseconds) of French narrative episodes*



The total duration of Italian recordings was 200 minutes; each audio recording consisted of free interaction in a naturalistic environment lasting ten minutes. Four of the dyads were eliminated from the analysis due to the acoustic quality of the recordings. 64 Italian narrative episodes (23.18 minutes in total) ( $21.73 \text{ sec} \pm 8.58$ ), ranging in milliseconds (8818.13 – 59155.30) were selected among the data of sixteen dyads (8 boys, 8 girls).

**Figure 4**

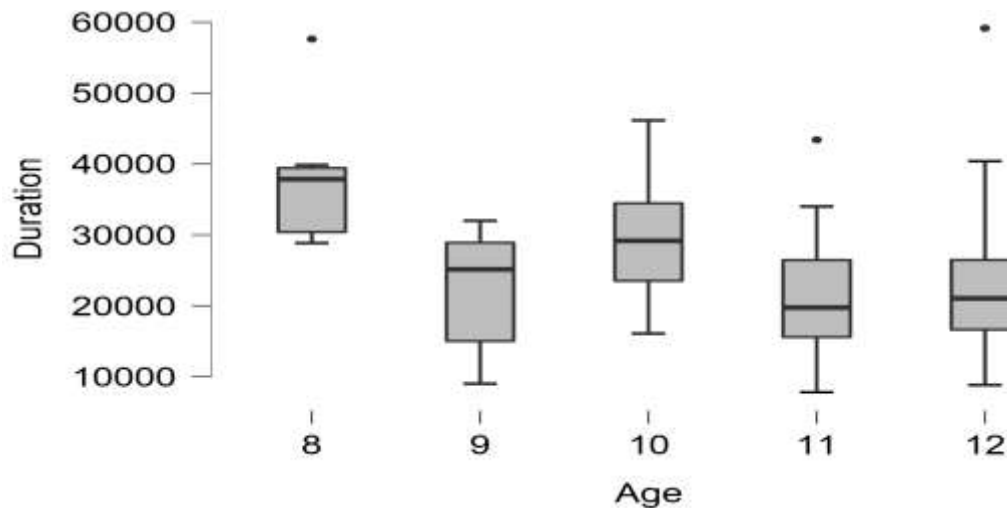
*Duration distribution (in milliseconds) of Italian narrative episodes*



Linear regression was conducted on the duration of narrative episodes. We observed a negative correlation between the age of infants ( $11.10 \pm 1.27$  in weeks) and duration ( $23.77 \pm 4.92$  s) of episodes  $R^2 = .11$ ,  $F(1,126) = 16.26$ ,  $p < .001$ ; while infant ages, duration of episodes decrease. Infant age was a controlled variable; however, there is an 8-to-12-week interval in French data, and an age effect appears.

**Figure 5**

*Duration (in milliseconds) distribution by age (in weeks)*



The box plot indicates that the duration of the episodes displays greater variability, especially for young-aged infants. Starting from 11 weeks old, a pattern seems to set in place. Still, at 12 weeks old, upper whiskers and outliers indicate signs of skewness.

Duration of narrative episodes between French ( $25.80 \text{ s} \pm 9.98$ ; range 7.81 – 57.63) and Italian ( $21.73 \text{ s} \pm 8.58$ ; range 8.82 – 59.16) is significantly different  $R^2 = .05$ ,  $F(1,126) = 6.12$ ,  $p = .01$ . No gender effect was observed ( $p = .55$ ).

### **6.1.2. Narrative Indicator Chart**

The narrative indicator chart includes common measures described in the narrative organization. The coding covers only acoustic measures, and no verbal content analysis is conducted. This analysis is performed based on the narrative indicators to contextualize the characteristics of French and Italian mother-infant dyads according to their respective interactive styles.

#### **Inter coder reliability for narrative indicator chart**

Two coders have coded 25% of data independently based on acoustic images and audio recordings. One episode corresponding to each mother-infant dyad was coded (16 French and 16 Italian episodes). Reliability between coders is calculated using Cronbach's alpha as an indicator ( $\alpha=0.74$ ).

## Analysis

Logistic regression is performed adjusted on age and gender to estimate the probability of indicators of a narrative. Odds ratios and Wald 95% Confidence Interval were calculated. A one-predictor logistic model was used to test the hypothesis regarding the relationship between the likelihood of a narrative indicator and nationality. A dichotomous outcome variable (Yes/No) from each indicator of the narrative chart is paired with a dichotomous predictor variable (nationality).

**Table 4**

*Proportions of narrative indicators for French and Italian dyads*

		French	Italian	Adj. p-value
<b>GLOBAL</b>				
	Bell-shaped pitch	40.63%	31.25%	0.27
<b>DYNAMIC</b>	Bell-shaped intensity	67.19%	78.13%	0.17
	Repeated sound	50%	54.69%	0.60
	Repeated words	37.50%	59.38%	0.01*
	Repeated group of words	70.31%	59.38%	0.20
	Repeated melodic contour	68.75%	78.13%	0.23
	Topic	100%	100%	-
	<b>BEGINNING</b>			
	Ascending contour	46.88%	46.88%	-
	Low intensity	85.94%	84.38%	0.80
	Pause before	67.19%	56.25%	0.20
	Topic change	37.50%	59.38%	0.01*
<b>MIDDLE</b>				
	Maximum frequency	26.56%	26.56%	-
	Maximum intensity	70.31%	75%	0.55
	Vocal turns	96.88%	100%	0.09
	Accelerated speech	79.69%	96.88%	<.01**
	Co-vocalizations	79.69%	96.88%	0.69
<b>ENDING</b>				
	Descending contour	60.94%	43.75%	0.06
	Slow speech rate	71.88%	85.94%	0.06
	Lengthening final	46.88%	39.06%	0.37
	Descending intensity	79.69%	87.50%	0.24
	Longer final pause	65.63%	57.81%	0.36

\*p < .05, \*\*p < .01, \*\*\*p < .001

For the global dynamic of the episodes, the logistic regression model was only significant for repeated words variable to help the prediction of indicating a narrative model. The model explained 6% of the variance (model Nagelkerke  $R^2 = 0.06$ ). The model correctly explains 61% of cases. Italian dyads are twice as likely to use repeated words (OR=2.44, 95%CI [1.20, 4.96],  $p < .05$ ).

Spearman correlations were conducted between indicators of global dynamics. The use of repeated words is highly correlated with repeated sound ( $r_s(126) = .46, p < .001$ ) and repeated melodic contour ( $r_s(126) = .30, p < .001$ ). Repeated sound and melodic contours are also positively correlated ( $r_s(126) = .28, p < .01$ ).

We have also conducted an analysis on the narrative indicator chart corresponding to the bell-shaped pitch and intensity curves. There is no correlation between intensity and fundamental frequency, which means the hypothesis suggesting the rise and fall of both intensity and fundamental frequency together is rejected ( $r_s(126) = -.08, p = .32$ ).

Topic change compared to previous utterances during the dialogue as a predictor increases the likelihood of predicting the model (model Nagelkerke  $R^2 = 0.06$ ); with Italians switching the topic more likely (OR=2.44, 95%CI [1.20, 4.96],  $p < .05$ ). No correlations have been found between indicators of the beginning part.

The logistic regression model is significant for accelerated speech rate, explaining 15% of the variance (model Nagelkerke  $R^2 = 0.15$ ). The validity of predicted probabilities is 88% of cases. Italian dyads are more likely to accelerate their speech (OR=7.90, 95%CI [1.70, 36.64],  $p < .01$ ). No correlations between items have been found.

For the ending part, significant correlations have been found for lengthening final, descending intensity ( $r_s(126) = .25, p < .01$ ) and longer final pause ( $r_s(126) = .26, p < .01$ ).

### **6.1.3. Principal Component Analysis**

We have conducted principal component analysis (PCA) on multiple indicators of the narrative chart to observe which dimensions appear prominent. PCA results give a new reference frame represented in dimensions. The narrative indicator chart contains four

major categories (global dynamic, beginning, middle, and ending) and a total of 21 items.

According to PCA, new dimensions appear to be the ones that have significant loadings. The first principal component (PC1) represents acoustic markers of the global fundamental frequency. Loadings show correlation with the markers of respective acoustic measures of the beginning (ascending contour), middle (maximum frequency value), and ending parts (descending contour), which are almost equally weighted. PC2 shows the same pattern as PC1 for the global intensity curve and is correlated with a respective acoustic measure of the beginning (low intensity), middle part (maximum intensity), and ending (descending intensity).

The first two dimensions, which are the most important principal components (PC1 & PC2) are representing 30% of the data. PCA revealed that five dimensions with an Eigenvalue  $>.1$  were considered significant, explaining 55% of the variance.

**Figure 6**

*Original variables and their relationship with principal components*

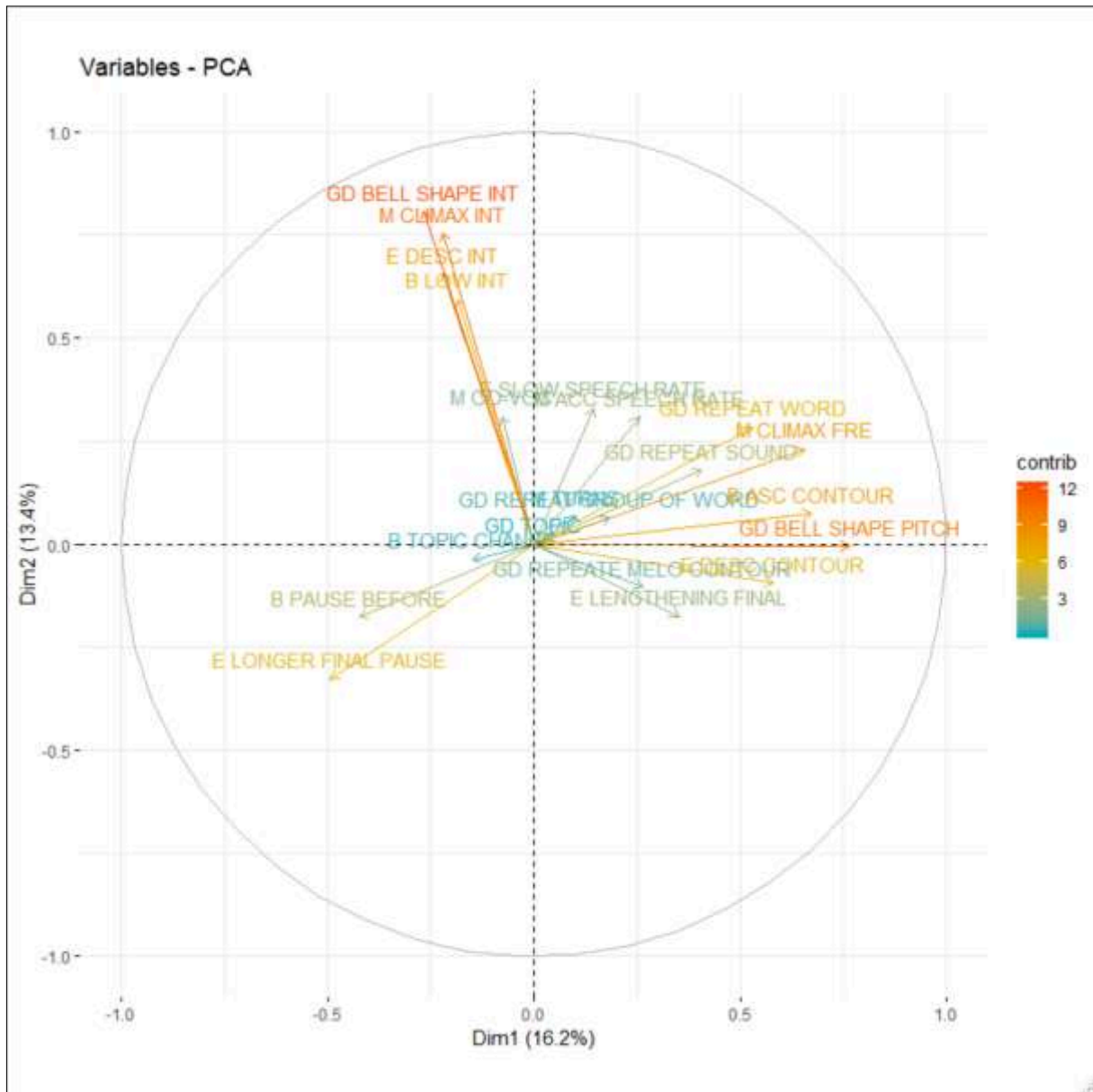


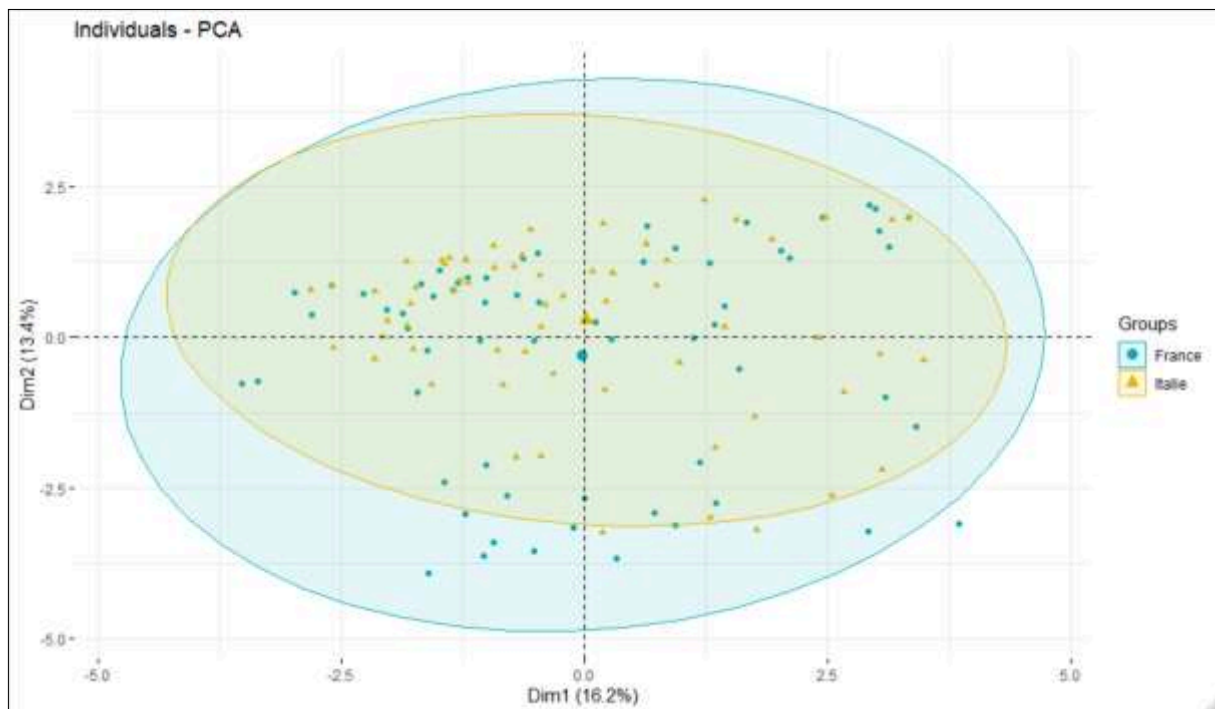
Figure 6 shows the correlation between the original variables and their relationship with two principal components. Acoustic measures of fundamental frequency and intensity appear as major contributors. PC3 represents repeated melodic contour items from the global dynamic (GD) category. Still, it has a small component in the direction of the co-vocalizations from the middle part (M), and PC4 represents the repeated group of words, again from the global dynamic (GD) category. Finally, PC5 represents the slow speech rate item from the ending (E) category. However, those



dimensions are not substantial contributors. PCA shows that items from the global dynamic category by themselves represent narrative patterning aspects; even so, it explains a low proportion of total variance.

**Figure 7**

*Distribution of original variables of French and Italian narratives*



PCA also demonstrates that French and Italian narratives show similar patterns, especially around Dimension 1 (fundamental frequency) and Dimension 2 (intensity). Notably, a cluster of intensity markers is exhibited in the figure above.

#### **6.1.4. Comparison of Vocal Units in French and Italian Narrative Episodes**

##### **6.1.4.1. French intra-group comparisons of narrative episodes based on vocal units**

The sample of French narrative episodes, from the total of 971 vocalizations, includes a total number of 783 maternal vocalizations, of which 491 are isolated, 150 overlapped that mother initiates, 140 overlapped that baby initiates and 2 overlapped in total sync (see Table 7). Maternal vocalizations are 50.7% of the total number of vocalizations.

Mother-only vocalizations lasted on average, 1455.98 ms  $\pm$  1038.01 (ranging from 732.41 to 3551.87).

The sample of French narratives includes a total number of 480 infant vocalizations, of which 188 are isolated, 150 overlapped that the mother initiates, 140 overlapped that baby initiates, and 2 overlapped in total sync. Contingency analysis revealed that no gender effect was found on the frequency of infant vocalizations between boys and girls (Chi square = 15.16;  $p = .06$ ; Cramer's  $V = 0.49$ ), nor on the frequency of maternal vocalizations (Chi square = 14.40;  $p = .42$ ; Cramer's  $V = 0.47$ ). No age effect was found on infant vocal frequency considering 8-to-12 week interval (Chi square = 28.39;  $p = .65$ ; Cramer's  $V = 0.33$ ). The number of maternal vocalizations increases while the infant ages (Chi square = 78.64;  $p = .02$ ; Cramer's  $V = 0.55$ ).

Infants' contribution is 31.1% of the total number of vocalizations. Isolated infant vocalizations lasted, on average, 915.30 ms  $\pm$  611.37 (ranging from 310.43 to 1848 ms).

571 pauses occurred within French narratives and lasted on average 908.65 ms  $\pm$  613.15 (ranging from 335.65 to 1687.31 ms), corresponding to 37% of the total number of vocalizations.

### Vocal rate

Pearson's correlation coefficients are calculated to examine intra-group vocal dynamics.

**Table 5**

*The vocal rate in French narrative episodes*

	<b>M (SD)</b>	<b>1</b>	<b>2</b>	<b>3</b>
1. Maternal rate	29.06 (9.38)	-		
2. Infant rate	18.84(12.57)	.73***	-	
3. Overlapped rate	11.15(11.04)	.86***	.90***	-
4. Pause rate	20.46(5.55)	-.08	-.14	-.25*

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Maternal vocal rate is strongly correlated with infant vocal rate and overlapped vocal rate. Infant vocal rate is also found to be strongly correlated with overlapped rates. Pause rates are only found to be weakly negatively correlated with overlapped rates.

**Table 6***Percentage of time spent vocalizing in French narrative episodes*

<b>%time</b>	<b>M (SD)</b>	<b>1</b>	<b>2</b>	<b>3</b>
1. Mother	.43(.16)	-		
2. Infant	.12(.10)	-.41***	-	
3. Overlapped	.15(.15)	-.63***	-.15	-
4. Pause	.30(.10)	-.20	-.09	-.40**

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ 

The percentage of time spent vocalizing displayed moderate negative correlations between infant and maternal vocalizations, maternal and overlapped vocalizations, and overlapped vocalizations and pauses in narrative episodes.

Duration of maternal vocalizations and pause duration are found to be moderately positively correlated ( $r(62) = .61, p < .001$ ). No correlation was found in terms of the durations of other types of vocalizations.

The number of pause occurrences is weakly positively correlated with both the number of infant vocalizations ( $r(62) = .33, p = .007$ ) and a strong positive correlation is found with the number of maternal vocalizations ( $r(62) = .82, p < .001$ ). The number of overlapped vocalizations that the infant initiates is strongly positively correlated with the ones that the mother initiates ( $r(62) = .96, p < .001$ ).

#### **6.1.4.2. Italian intra-group comparisons of narrative episodes based on vocal units**

The sample of Italian narratives includes a total number of 493 maternal vocalizations, 82 overlapped that mother initiates, 103 overlapped that baby initiates and 1 overlapped in total sync, corresponding to 55.5% of the total number of vocalizations. Isolated maternal vocalizations lasted on average  $1550.53 \text{ ms} \pm 1224.96$ , ranging from 943.15 to 3003.94. Duration of vocalizations and pauses of Italian sample show only correlations between maternal and infant vocalizations  $r(62) = .35, p < .01$ ; and between maternal vocalizations and pause durations  $r(62) = .48, p < .001$ . The number of maternal vocalizations is also significantly correlated with the number of infant vocalizations  $r(62) = .58, p < .001$ , and the number of pauses  $r(62) = .84, p < .001$ .

The sample of Italian narratives includes a total number of 146 isolated infant vocalizations, 82 overlapped that mother initiates, 103 overlapped that baby initiates, and 1 overlapped in total sync, corresponding to 27.1% of total vocalizations. Infant-only vocalizations lasted on average  $727.93 \pm 423.27$ , ranging from 179.25 to 2333.6. 399 pauses occurred in Italian data, corresponding to 32.6% of total vocalizations. Pauses lasted on average  $823.26 \text{ ms} \pm 458.53$ , ranging from 431.47 to 2322.50. No gender effect was found on infant vocalizations ( $p = .24$ ) nor on maternal vocalizations ( $p = .13$ ) in the Italian group.

Infant vocalizations are equally distributed in both French ( $n=188$ , 12.2%) and Italian ( $n=146$ , 11.9%) groups. Maternal vocalizations are also equally distributed in terms of numbers for French ( $n=491$ ) and Italians ( $n=493$ ). However, French mothers' vocalizations constitute 31.8% of the total number of vocalizations, while Italian mothers' vocalizations are 40.3%. More within and between-speaker pauses occurred in the French group (37% vs. 32.6%).

**Table 7**

*Number of vocal occurrences in French and Italian narrative episodes*

Vocalizations	Italian N=64		French N=64	
	Number	%	Number	%
Baby	146	11.9	188	12.2
Overlapped Baby initiates	103	8.4	140	9.1
Mother	493	40.3	491	31.8
Overlapped Mother initiates	82	6.7	150	9.7
Overlapped	1	0.1	2	0.1
Pauses	399	32.6	571	37
Total	1224	100	1542	100

### Vocal rate

Pearson's correlation coefficients are calculated to examine Italian intra-group vocal dynamics.

**Table 8***Vocal rates in Italian episodes*

Rate	M (SD)	1	2	3
1. Mother	29.7(7.33)	-		
2. Infant	14.15(9.21)	.73***	-	
3. Overlapped	8.04(8.25)	.75***	.89***	-
4. Pause	17.39(5.0)	.10	-.36**	-.30*

\* p < .05, \*\* p < .01, \*\*\* p < .001

Maternal and infant vocal rates show a significant positive correlation as well as overlapped vocal rates. Pause rates are negatively correlated with both infant and overlapped vocal rates.

**Table 9***Percentage of time spent vocalizing in Italian dyads*

%time	M (SD)	1	2	3
1. Mother	.55(.13)	-		
2. Infant	.07(.06)	-.23	-	
3. Overlapped	.13(.14)	-.68***	-.19	-
4. Pause	.24(.09)	-.25	-.04	-.42***

\* p < .05, \*\* p < .01, \*\*\* p < .001

**6.1.4.3. Comparison of vocal units between French and Italian narrative episodes**

Multiple regressions, adjusted on age and gender, were conducted to treat the nationality variable on different types of vocalizations along with the measure of effect size.

An additional regression model, including age and gender, is also used to elaborate on the results, and the standardized regression coefficient is calculated.

## Duration of vocalizations

**Table 10**

*Comparison of durations of vocal units (in milliseconds)*

Duration (ms)	Italian N=64	French N=64	$\eta^2_p$	<i>p</i> -value
Mother	11944 (5692.85)	11170.16 (6042.08)	-	NS
Infant	1660.61 (1529.88)	2688.69 (2355.81)	.04	.02
Overlapped	2996.57 (3444.75)	3837.17 (4270.96)	-	NS
Pauses	5132.56 (2369.20)	8106.93 (4403.27)	-	NS

On the duration of maternal vocalizations, no significant difference was found between French and Italian (11170.16±6042.08 ms) vs. 11944±5692.85 ms) mothers ( $p = .46$ ). Neither infant age ( $p = .10$ ) nor gender of the infant ( $p = .42$ ) affects maternal vocal durations. Duration of overlapped vocalizations (French, 3837.17±4270.96 ms) vs. Italian, 2996.57±3444.75 ms) display no significant difference ( $p = .22$ ).

**Table 11**

*Regression models for gender and nationality effects on infant vocal duration*

Variable	$R^2$ Change	<i>F</i> Change	$\beta$	<i>p</i>	95%CI	
					Lower	Upper
Model 1						
Gender	.06	7.92	.26	< .001	373.92	1741.65
Model 2						
Nationality	.07	10.02	-.26	.02	-1751.36	-404

The overall model predicted approximately 13% of the variance in infant vocal durations,  $R^2 = .13$ ,  $F(2,125) = 9.26$ ,  $p < .001$ . The first model, including gender effects, predicts approximately 6% of the variance in the duration of vocalizations and indicates girls vocalize longer than boys ( $R^2 = .06$ ,  $F(1,126) = 7.92$ ,  $p < .001$ ). The second model, including the nationality variable, predicts an additional 7% of the total variance and

indicates French babies (2688.69 ms ± 2355.81) vocalize longer than Italian babies (1660.61 ms ± 1529.88). No age effect was found ( $p = .85$ ).

**Table 12**

*Regression models for age and nationality effects on pause durations*

Variable	$R^2$ Change	F Change	$\beta$	p	95%CI	
					Lower	Upper
Model 1						
Age	.27	45.70	-.48	< .001	-2110.43	-806.35
Model 2						
Nationality	.00	0.18	-.05	.67	-1998.87	1291.19

The overall model predicted approximately 27% of the variance in pause durations,  $R^2 = .27$ ,  $F(2,125) = 22.79$ ,  $p < .001$ . The regression model displays that variance comes from the 8 to 12 week age interval in French data. The nationality variable has no effect on the duration of pauses. No gender effect was found on the duration of pauses ( $p = .47$ ).

### Vocal rate

**Table 13**

*Comparison of vocal rates between French and Italian dyads*

Vocal Rate	Italian N=64	French N=64	$\eta^2_p$	p-value
Mother	29.7 (7.33)	29.1 (9.4)	-	NS
Infant	14.15 (9.21)	18.8 (12.5)	.07	.02
Overlapped	8.04 (8.25)	11.15 (11.04)	.03	.07
Pauses	17.39 (5)	20.5 (5.5)	-	NS

Maternal vocal rate shows no difference between two groups ( $p = .67$ ). French babies' vocal rate is higher than Italian babies  $R^2 = .04$ ,  $F(1,126) = 5.79$ ,  $p = .02$ . Gender ( $p = .36$ ) or age ( $p = .09$ ) have no effect on infant vocal rate. Overlapped vocal rate is marginally high in French dyads. Pause rate does not show any difference between the

two nationalities ( $p = .27$ ), including gender ( $p = .68$ ); even so, a marginal age effect is detected ( $p = .08$ ).

### Percentage of time spent vocalizing

**Table 14**

*Comparison of percentage time spent vocalizing*

Time%	Italian N=64	French N=64	$\eta^2_p$	$p$ -value
Mother	.55 (.13)	.43 (.16)	.16	< .001
Infant	.07 (.06)	.12 (.10)	.07	< .001
Overlapped	.13 (.14)	.15 (.15)	-	NS
Pauses	.24 (.09)	.30 (.10)	-	NS

Italian mothers spent more time vocalizing than French mothers  $R^2 = .16$ ,  $F(1,126) = 23.29$ ,  $p < .001$ . French babies have more time percent vocalized in episodes than Italian babies  $R^2 = .07$ ,  $F(1,126) = 8.86$ ,  $p < .001$ ; moreover, no gender ( $p = .10$ ) or age ( $p = .35$ ) effects were observed. No difference was found between pause rates based on nationality; however, age interval among French infants explains a small amount of variance between groups ( $\eta^2_p = .05$ ,  $p = .01$ ).

### Number of vocalizations

**Table 15**

*Comparison of numbers of vocalizations*

Number	Italian N=64	French N=64	$\eta^2_p$	$p$ -value
Mother	7.70 (3.17)	7.67 (3.84)	-	NS
Infant	2.28 (1.93)	2.94 (2.01)	-	NS
Overlapped	2.91 (3.17)	4.56 (5.08)	.04	.03
Pauses	6.23 (2.86)	8.92 (4.28)	-	NS

There is no difference between French and Italian dyads regarding the number of maternal ( $p = .96$ ) or infant ( $p = .26$ ) vocalizations. However, a gender effect on the number of vocalizations was observed, showing that girls vocalize more than boys ( $3.11 \pm 2.13$  vs.  $1.91 \pm 1.52$ ,  $\eta^2_p = .10$ ,  $p < .001$ ). An age effect is observed concerning the number of pauses with a decrease while the infant ages ( $\eta^2_p = .11$ ,  $p < .001$ ).



#### 6.1.4.4. Comparison of acoustic measures between French and Italian narrative episodes

We have conducted independent samples t-test to see inter-group differences between individual vocalizations.

**Table 16**

*Comparison of the fundamental frequency of different types of vocalizations*

<b>Fundamental Frequency (Hz)</b>	<b>Infant</b>	<b>Mother</b>	<b>Overlapped</b>
French	2173.87(1045.21)	3559.21(1331.50)	2674.55(1317.95)
Italian	1974.53(691.04)	3016.95(1174.21)	2201.26(932.68)

On frequency values, French babies display high-frequency values than Italians,  $t(332) = 2.504, p = .01$ . French mothers show higher frequency values,  $t(982) = 7.06, p < .001$  than Italian mothers. Frequency values of overlapped vocalizations are significantly higher for French dyads,  $t(474) = 4.26, p < .001$ .

To examine the internal acoustic variability of narrative episodes, further analysis is made on the overall range of fundamental frequency of French and Italian episodes adjusted on gender and age.

**Table 17**

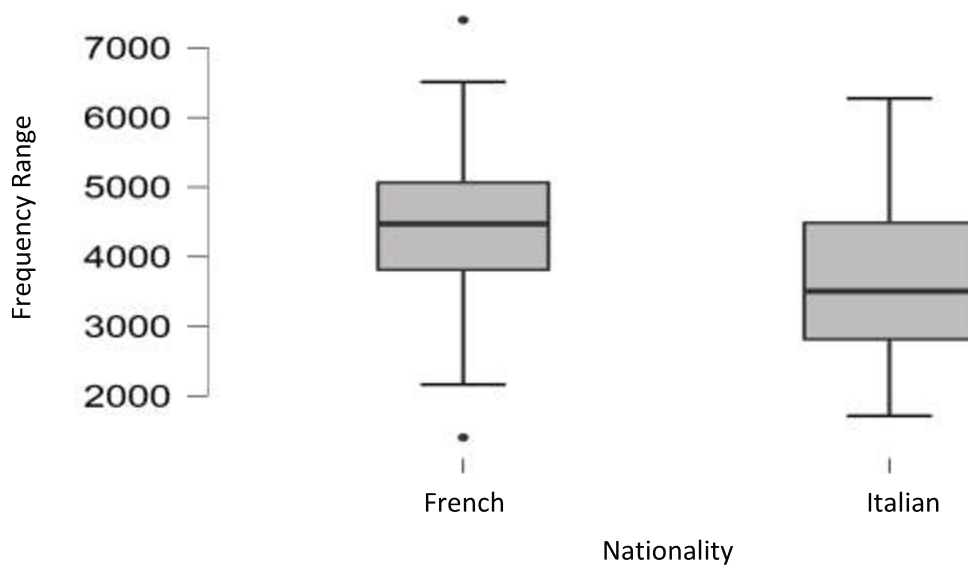
*Overall Frequency Range of French and Italian Narratives*

<b>Fundamental Frequency (Hz)</b>	<b>Mean</b>	<b>SD</b>	<b>Range</b>
French	4463.51	1056.79	1404.36 – 7401.67
Italian	3611	1074.15	1713.29 – 6275

French narratives ( $4463.51 \pm 1056.78$  Hz; ranged 1404.36 - 7401.66) is significantly different from Italian narratives ( $3610.99 \pm 1074.15$  Hz; ranged 1713.28 - 6275) based on overall frequency range  $R^2 = .17, F(3,124) = 8.39, p < .001$ .

**Figure 8**

*Frequency Range of French and Italian Narrative Episodes*



Italian data seems to have larger frequency ranges indicating a wider distribution. Italian frequency distribution is positively skewed. French frequency range seems to be symmetrically distributed; however, upper and lower outliers are demonstrated.

The range of intensity of French narratives (28.65 dB  $\pm$  4.51; ranged 17.77-42.46) is also significantly different from Italian narratives (24.14 dB  $\pm$  4.15; ranged 15.66-38.84) based on the overall range of episodes  $R^2 = .26$ ,  $F(3,124) = 14.55$ ,  $p < .001$ .

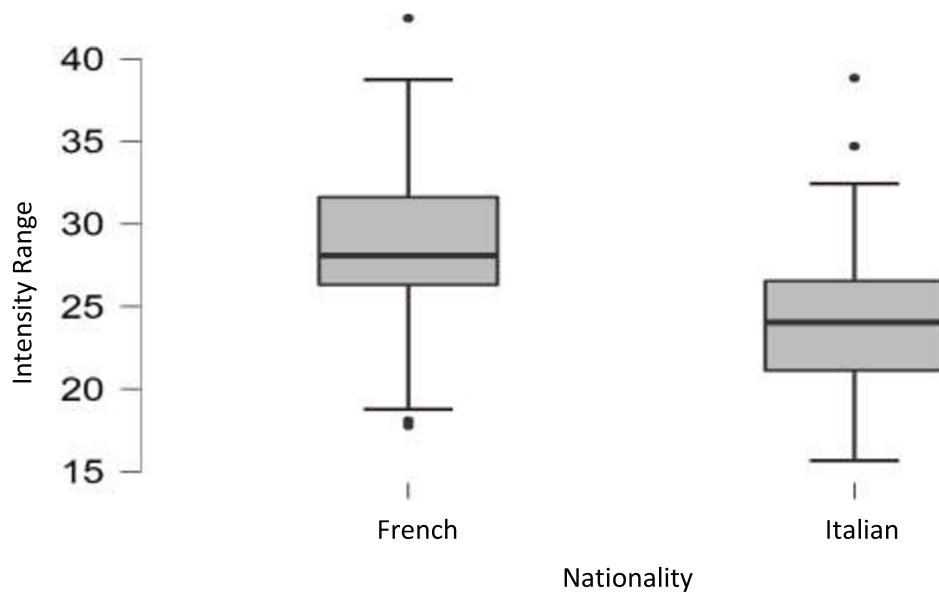
**Table 18**

*Overall intensity range of French and Italian narrative episodes*

Intensity (dB)	Mean	SD	Range
French	28.65	4.52	17.77 – 42.46
Italian	24.14	4.15	15.66 – 38.84

**Figure 9**

*Intensity range of French and Italian Narrative Episodes*



The box plot displays that there is a greater variability regarding the intensity range of French data, as well as outliers. The interquartile range of Italian data is equally distributed around the median. However, greater outliers are also demonstrated.

## **6.2. Study 2: Comparison of vocal units in French narrative and control episodes**

The sample of French control episodes includes a total number of 419 isolated maternal vocalizations: 17 overlapped that mother initiates, 8 overlapped that baby initiates, and 19 overlapped in total sync (see Table 20). Maternal vocalizations make up 46.6% of the total number of vocalizations. Mother-only vocalizations lasted, on average, 1471.57 ms  $\pm$  1230.58 (ranging from 614.68 to 2782.36).

The sample of French control episodes includes a total number of 101 isolated infant vocalizations, 17 overlapped that mother initiates, 8 overlapped that baby initiates, and 19 overlapped in total sync, making 14.6% of the total number of vocalizations. The number of infant vocalizations showed no age effect (Chi-square = 34.45;  $p = .26$ ; Cramer's  $V = 0.34$ ). The number of maternal vocalizations does not differ while the infant ages (Chi-square = 61.91;  $p = .59$ ; Cramer's  $V = 0.45$ ). Infant-only vocalizations lasted, on average, 811.82 ms  $\pm$  527.25 (ranging from 335.80 to 2542.30). 429 pauses

occurred in control episodes, corresponding to 43.2% of the total number of vocalizations. Pauses lasted  $1517.15 \text{ ms} \pm 1432.98$  and ranged from 636.86 to 4267.93.

**Table 19**

*Vocal units in French data*

Vocalizations	Control N=60		Narrative N=64	
	Number	%	Number	%
Baby	101	10.2	188	12.2
Overlapped Baby initiates	8	0.8	140	9.1
Mother	419	42.2	491	31.8
Overlapped Mother initiates	17	1.7	150	9.7
Overlapped	19	1.9	2	0.1
Pauses	429	43.2	571	37
Total	993	100	1542	100

### 6.2.1. French intra-group comparisons of control episodes based on vocal units

Within control episodes, the maternal vocal rate is correlated with the pause rate ( $r(58) = .79, p < .001$ ). Infant vocal rate is correlated with the overlapped rate ( $r(58) = .61, p < .001$ ).

Infant vocal durations are correlated with overlapped durations ( $r(58) = .33, p = .009$ ). Maternal vocal durations are correlated with pause durations ( $r(58) = .32, p = .01$ ).

Pause percentage is negatively correlated with mothers' ( $r(58) = -.78, p < .001$ ) and overlapped percentage ( $r(58) = -.52, p < .001$ ).

The number of occurrences of infant and maternal vocalizations is found to be weakly correlated ( $r(58) = .27, p = .04$ ). Number of maternal vocalizations and pauses are strongly correlated ( $r(58) = .90, p < .001$ ), along with a moderate correlation between the number of infant vocalizations and pauses ( $r(58) = .47, p < .001$ ).

### 6.2.2. Comparison of vocal units in narrative and control episodes within French data

Multiple regressions were conducted to compute partial correlation coefficients to calculate effect sizes.

### ***Duration of vocalizations***

**Table 20**

*Comparison of durations of different types of vocalizations and pauses*

<b>Duration (ms)</b>	<b>Control N=60</b>	<b>Narrative N=64</b>	<b><math>\eta^2_p</math></b>	<b><i>p</i>-value</b>
Mother	10276.50 (5705.32)	11170.16 (6042.08)	-	NS
Infant	1366.56 (1338.61)	2688.69 (2355.81)	.11	< .001
Overlapped	1253.57 (2034.29)	3837.17 (4270.96)	.13	< .001
Pauses	10847.66 (5287.23)	8106.93 (4403.27)	.08	.002

Duration of maternal vocalizations does not differ in narrative and control episodes ( $p = .39$ ). Infants vocalize significantly longer in narrative episodes, with partial eta square showing that the proportion of variance between groups on the set of outcomes is .11, indicating a small effect.

**Table 21**

*Regression models for gender and narrative effects on infant vocal duration*

<i>Variable</i>	<i>R<sup>2</sup> Change</i>	<i>F Change</i>	$\beta$	<i>p</i>	<i>95%CI</i>	
					<i>Lower</i>	<i>Upper</i>
<b>Model 1</b>						
Gender	.9	11.38	.28	< .001	490.43	1813.66
<b>Model 2</b>						
Narrative	.11	15.0	.32	< .001	629.68	1947.38

The overall model predicted approximately 19% of the variance in infant vocal durations,  $R^2 = .19$ ,  $F(2,121) = 13.84$ ,  $p < .001$ . The first model, including gender effects, predicts approximately 9% of the variance in the duration of vocalizations and indicates girls vocalize longer than boys ( $R^2 = .09$ ,  $F(1,122) = 11.38$ ,  $p < .001$ ). The second model, including the narrative effect, predicts an additional 11% of the total variance.

Overlapped vocalizations in narrative episodes last longer than control episodes  $R^2 = .14$ ,  $F(2,121) = 9.46$ ,  $p < .001$ . No gender effect was found ( $p = .36$ ).

Pause durations last longer in control episodes than in narratives  $R^2 = .08$ ,  $F(1,122) = 9.89$ ,  $p = .002$ . Furthermore, pause durations ( $1100.27 \pm 1015.75$  for control episodes vs.  $552.63 \pm 229.98$  for narratives) are more variable in control episodes than narratives  $r = -0.35$  ( $t(122) = -4.17$ ;  $p < .001$ ,  $R^2 = 0.13$ ).

### **Vocal rates**

**Table 22**

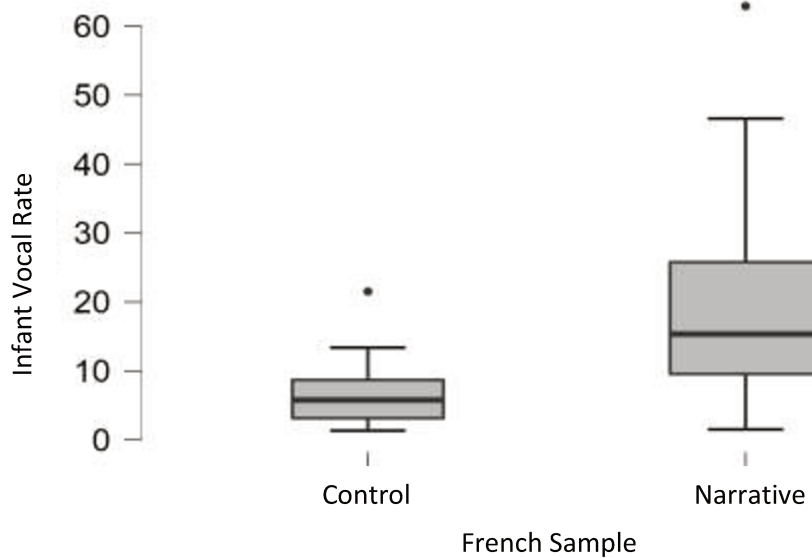
*Comparison of vocal rates between French narrative and control episodes*

<b>Vocal Rate</b>	<b>Control N=60</b>	<b>Narrative N=64</b>	<b><math>\eta^2_p</math></b>	<b>p-value</b>
Mother	20.6 (6.0)	29.1 (9.4)	.23	< .001
Infant	6.3 (3.7)	18.8 (12.5)	.31	< .001
Overlapped	1.90 (2.60)	11.15 (11.04)	.25	< .001
Pauses	18.74 (5.0)	20.5 (5.5)	.03	.07

Pause rates show no difference in narrative and control episodes. Maternal vocal rate is found to be significantly higher in narrative episodes compared to control episodes ( $R^2 = .23$ ,  $F(1,122) = 35.58$ ,  $p < .001$ ), as well as overlapped vocal rates ( $R^2 = .25$ ,  $F(2,121) = 20.35$ ,  $p < .001$ ), no gender effect is observed ( $p = .39$ ). Gender of the infant has no effect on maternal vocal rate ( $p = .22$ ). Infant vocal rate in narrative episodes are also significantly higher than in control episodes; the overall model predicted approximately 32% of variance in infant vocal rates,  $R^2 = .32$ ,  $F(2,121) = 28.08$ ,  $p < .001$ . No gender effect is found ( $p = .27$ ). The values of partial eta square for each vocal rate, shown in Table 2, suggest that narrative effects for infant ( $\eta^2_p = .31$ ), maternal ( $\eta^2_p = .23$ ), and overlapped rate ( $\eta^2_p = .25$ ) are strong.

**Figure 10**

*Infant vocal rate in French narrative and control episodes*



Apparent differences are represented in the box plots diagram. Infant vocal rate in narrative episodes displays more significant variability than in control episodes, in which infant vocal rate shows a compact distribution around the median.

***Percentage of time spent vocalizing***

We calculated the percentage of time vocalized within each episode across narrative and control modalities.

**Table 23**

*Comparison of time % spent vocalizing in French narrative and control episodes*

Time%	Control N=60	Narrative N=64	$\eta^2_p$	<i>p</i> value
Mother	.43 (.13)	.43 (.16)	-	NS
Infant	.06 (.05)	.12 (.10)	.12	< .001
Overlapped	.05 (.07)	.15 (.15)	.15	< .001
Pauses	.50 (.14)	.30 (.10)	.30	< .001

Mothers spent a similar amount of time vocalizing in narrative and control episodes ( $p= .99$ ). Percentage of time spent in overlapped vocalizations in narrative episodes is

higher than in control episodes ( $R^2 = .17$ ,  $F(2,121) = 12.34$ ,  $p < .001$ ), showing a large effect size ( $\eta^2_p = .15$ ), but no gender effect was observed ( $p = .19$ ). Pauses in control episodes have a greater percentage of time than in narrative episodes ( $R^2 = .30$ ,  $F(1,122) = 52.13$ ,  $p < .001$ ) and show a strong effect ( $\eta^2_p = .30$ ). Infants spend more significant percentage of time vocalized in narrative episodes ( $R^2 = .23$ ,  $F(2,121) = 18.34$ ,  $p < .001$ ), also show a significant gender effect with girls greater than boys ( $R^2 = .12$ ,  $F(1,122) = 16.13$ ,  $p < .001$ ).

### ***The number of vocalizations***

**Table 24**

*Comparison of the number of occurrences of vocalizations*

<b>Number</b>	<b>Control N=60</b>	<b>Narrative N=64</b>	<b><math>\eta^2_p</math></b>	<b><i>p</i>-value</b>
Mother	6.98 (3.01)	7.67 (3.84)	-	NS
Infant	1.68 (1.35)	2.94 (2.01)	.12	< .001
Overlapped	0.73 (1.01)	4.56 (5.08)	.21	< .001
Pauses	7.15 (3.07)	8.92 (4.28)	-	NS

Number of maternal vocalizations shows no difference between narrative and control episodes ( $p = .27$ ). Infants vocalize more in narrative episodes than control episodes ( $R^2 = .20$ ,  $F(2,121) = 15.36$ ,  $p < .001$ ), showing also a gender effect revealing that girls vocalize more than boys ( $p < .001$ ), explaining approximately 9% of the total variance.

More overlapped vocalizations were observed in narrative episodes compared to control episodes ( $R^2 = .22$ ,  $F(2,121) = 16.80$ ,  $p < .001$ ), with a large effect size, but no gender difference was found ( $p = .37$ ). No difference was found between the number of pauses in narrative and control episodes.

### **6.2.3. Comparison of acoustic measures between French narrative and control episodes**

Overall mean fundamental frequency and intensity values of French narrative and non-narrative episodes are calculated to compare the density of their respective acoustic values.



**Table 25**

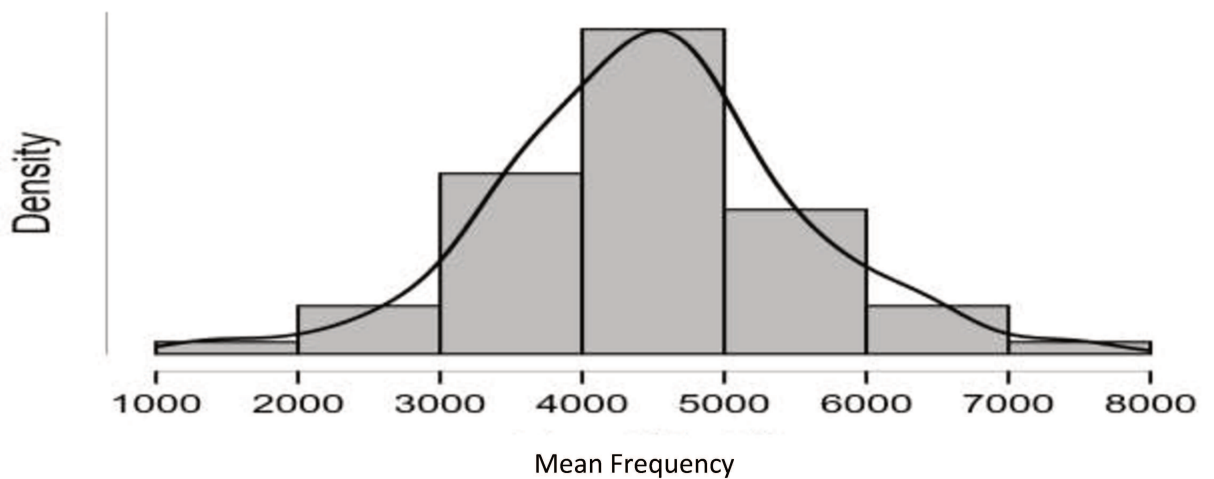
*Measures of fundamental frequency in French narrative and control episodes*

<b>Fundamental Frequency (H)</b>	<b>Mean</b>	<b>SD</b>	<b>Range</b>
Narrative	4463.51	1056.79	1404.36 – 7401.67
Control	4643.09	1169.57	1685.80 – 7233.11

Narrative and control episodes do not differ significantly regarding fundamental frequency ( $p = .37$ ).

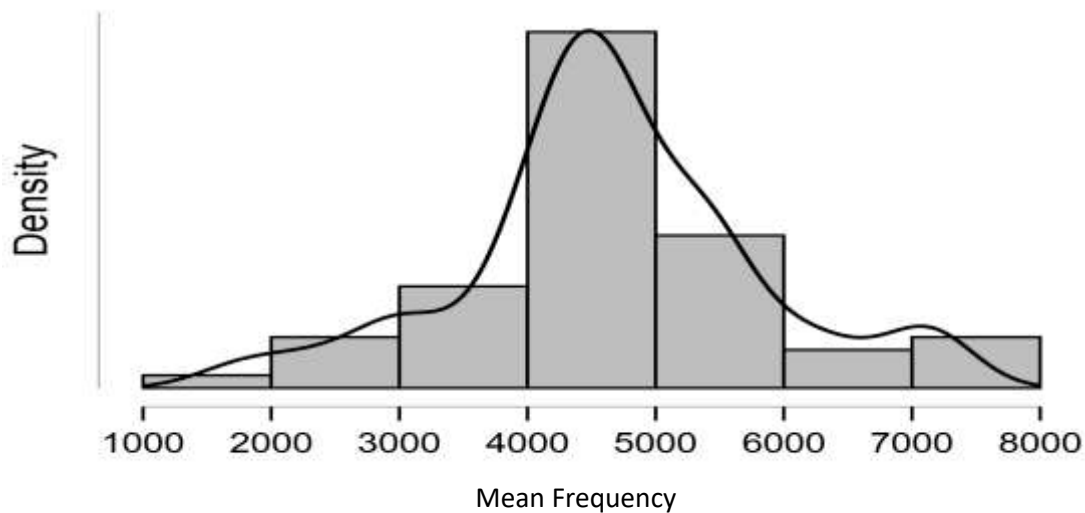
**Figure 11**

*Distribution of fundamental frequency in French narrative episodes*



**Figure 12**

*Distribution of fundamental frequency in French control episodes*



**Table 26**

*Measures of the fundamental frequency of different types of vocalizations*

<b>Fundamental Frequency (Hz)</b>	<b>Infant</b>	<b>Mother</b>	<b>Overlapped</b>
Narrative	2173.87(1045.21)	3559.21(1331.50)	2674.08(1305.18)
Control	1845.30(1079)	3481.90(1395.34)	2894.20(1340.46)

We have also calculated the frequency range and mean values of infant and maternal vocalizations for each paired episode between groups. The paired samples test reveals that both frequency ranges ( $t(59) = 2.96, p = .004$ ) and mean frequency values ( $t(59) = 3.57, p = .001$ ) of infant vocalizations are significantly higher in narrative episodes. The mean frequency of maternal vocalizations is also higher in narrative episodes ( $t(59) = 2.11, p = .04$ ). No difference was found between overlapped frequency values ( $p = .69$ ).

**Table 27**

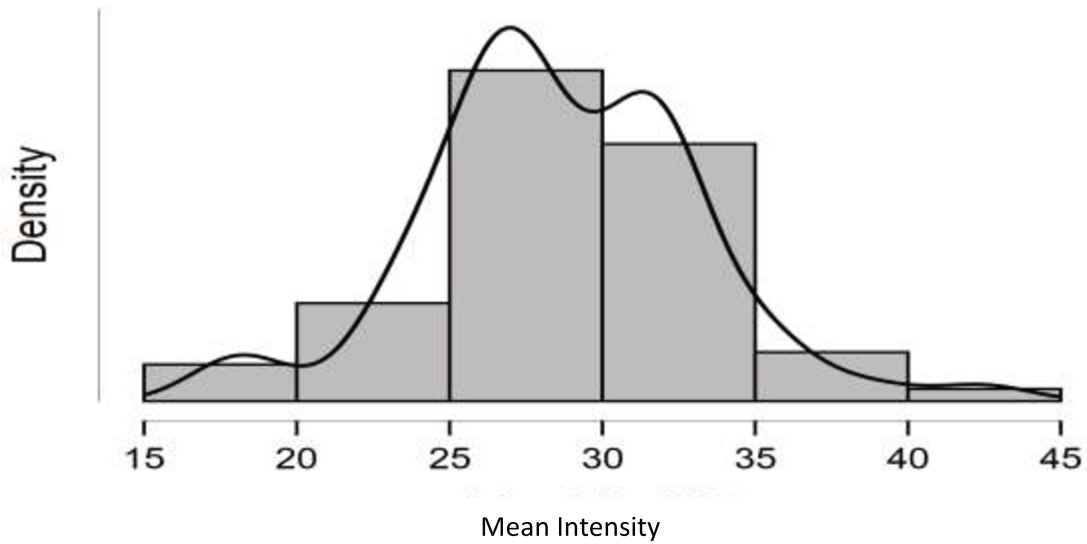
*Measures of intensity in French narrative and control episodes*

<b>Intensity (dB)</b>	<b>Mean</b>	<b>SD</b>	<b>Range</b>
Narrative	28.65	4.52	17.77 – 42.46
Control	25.79	5.68	12.52 – 39.03

Overall intensity is significantly higher in narrative episodes compared to control episodes  $R^2 = .07$ ,  $F(1,122) = 9.72$ ,  $p = .002$ ).

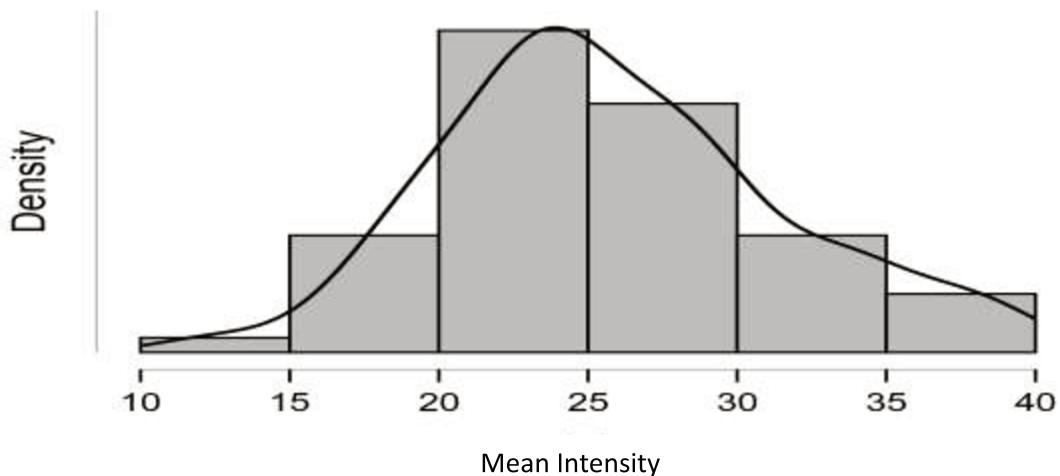
**Figure 13**

*Distribution of intensity in French narrative episodes*



**Figure 14**

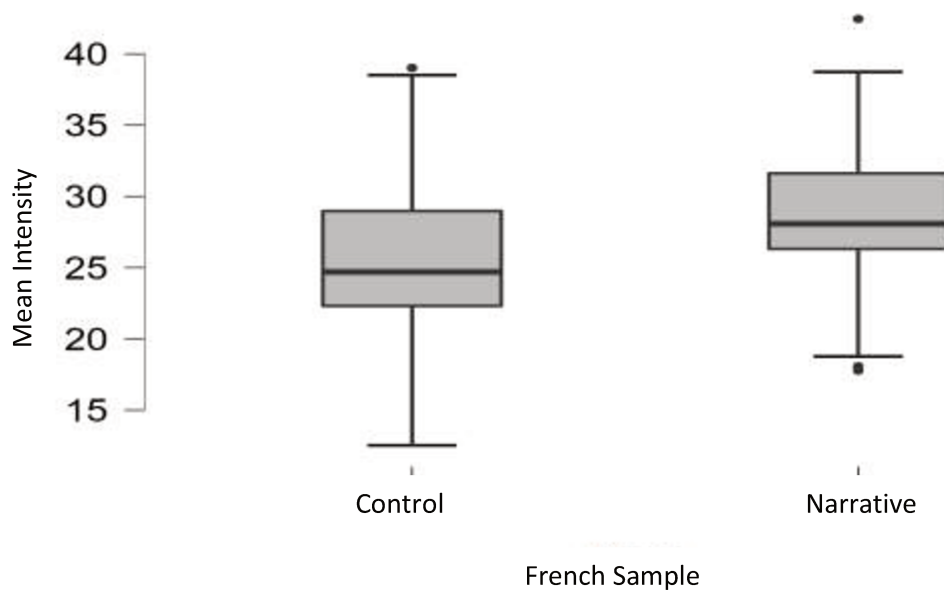
*Distribution of intensity in French control episodes*



The Control group intensity range stretches over a broader range of values around the median, while the narrative group intensity appears to have a more symmetrical distribution.

**Figure 15**

*Intensity range in French narrative and control episodes*



**Table 28**

*Measures of the intensity of different types of vocalizations*

Intensity (dB)	Infant	Mother	Overlapped
Narrative	27.92(5.60)	25.27(5.20)	28.53(5.96)
Control	26.03(5.60)	23.75(5.73)	27.86(4.98)

Paired samples test shows that intensity range of infant vocalizations is more variable in narrative episodes ( $t(59) = 2.34, p = .02$ ). Mean intensity values are also higher in narrative episodes for infant vocalizations ( $t(59) = 2.37, p = .001$ ).

Maternal vocalizations only differ in terms of variability of intensity ( $t(59) = 2.79, p = .007$ ). No difference was found between overlapped intensity values ( $p = .95$ ).

#### **6.2.4. Narrative analysis of verbal content in narrative and control episodes**

The verbal content analysis is performed on French narrative and control episodes. We anticipate significant differences according to indicators specific to the narrative organization. For example, more variations of tense, time, and temporal order are expected to be observed in narrative episodes compared to control episodes. However, no significant difference is expected for the indicators shared with infant-directed speech, such as repetition and use of questions.

#### **Inter-coder reliability for verbal content**

Inter-coder reliability was tested on the verbal content of the French narrative episodes and the control group on context, chronology, theme, and agency. The Cronbach alpha reliability estimates a consistency between coders ( $\alpha=0.72$ ).

#### **Analysis**

We have conducted a logistic regression, adjusted on age and gender, to compare verbal content items between narrative and control episodes of French data.

In the context category, tense use, the dominance of present tense use, and time and place indicators were examined. However, no significant difference was found between narrative and control episodes except for words indicating time (OR=4.38, 95% CI [1.17, 16.41]).

**Table 29***Verbal content analysis of French narrative and control episodes*

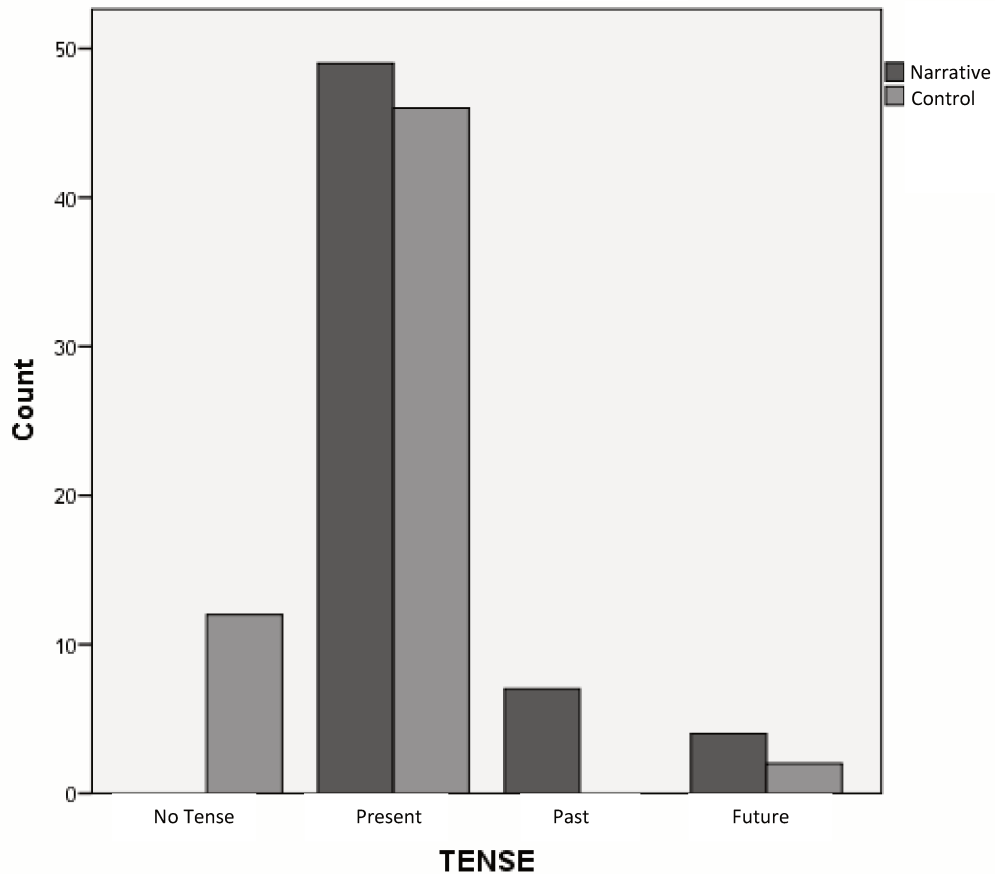
	Narrative	Control	Adj. p value
<b>CONTEXT</b>			
Tense	100%	80%	0.99
Present	81%	77%	0.53
Place	4%	0%	0.99
Time	18%	5%	0.02*
<b>THEME</b>			
Adjectives	61%	18%	<.001***
Questions	69%	60%	0.31
Recurrent Words	84%	68%	0.03*
Interpretation	78%	25%	<.001***
<b>CHRONOLOGY</b>			
Repetition	83%	75%	0.29
Elaboration	59%	13%	<.001***
End marker	21%	2%	0.008**
Temporal Order	25%	5%	0.28
<b>AGENCY</b>			
You	94%	75%	0.007**
Me	19%	7%	0.05*
Us	16%	7%	0.12
Emotion Attribution	27%	8%	0.01*
Cognition Attribution	52%	5%	<.001***
Appellatives	25%	5%	0.005**

\* p < .05, \*\* p < .01, \*\*\* p < .001

A Chi-square test was conducted to see the use of tenses in narratives and the control group ( $\chi^2 (3) = 27.11; p < .001; Cramer's V=0.41$ ). Figure 16 represents the distribution of tense use. No tense category is only observed within the control group ( $n=12, 20\%$ ). Use of present tense is equally distributed between groups ( $n =46, 81.7\%$  for narratives vs.  $n = 49, 76.7\%$  for control) with overall use of 79.2%. The past tense was used only in the narrative group ( $n= 7, 11.7\%$ ). The future tense is observed in both groups, with a low proportion of 6.7% for narratives ( $n = 4$ ) and 3.3% for control ( $n = 2$ ).

**Figure 16**

*Tense use in narrative and control episodes*



In the theme category, the use of adjectives (OR = 6.95, 95%CI [3.05, 15.85]) and interpretation of infant behavior (OR = 10.71, 95%CI [4.66, 24.63]) occurring within narrative episodes were higher than in control episodes. Elaboration is significantly determining the chronological flow within the narrative episodes (OR = 9.50, 95%CI [3.88, 23.27]). End markers are also highly used in narratives (OR = 16.52, 95%CI [2.10, 130.01]). However, this item does not provide a precise representation due to a large confidence interval.

Mothers use the “you” pronoun in both narrative and control episodes; however, appellatives are mostly used in narratives (OR = 6.33, 95%CI [1.74, 23.04]). Attribution

of emotion (OR = 3.98, 95%CI [1.36, 11.60]) and cognition (OR = 20.23, 95%CI [5.74, 71.32]) have a higher place within narrative episodes.

To test further hypotheses within narratives' verbal content, we have conducted correlations on some significant markers:

**Table 30**

*Correlations between verbal content indicators*

Variable	1	2	3	4	5	6	7	8	9	10
1. REPETITION	—									
2.ELABORATION	0.19*	—								
3.INTERPRETATION	0.02	0.43***	—							
4.EMOTION	-0.02	0.26**	0.36***	—						
5.ME	0.02	0.1	0.17	0.07	—					
6.YOU	0.11	0.19*	0.13	0.08	-0.04	—				
7.COGNITION	-0.02	0.32***	0.4***	0.07	0.07	0.22*	—			
8.ADJECTIVES	0.1	0.42***	0.19*	0.05	0.03	0.21*	0.24**	—		
9.END MARKERS	0.01	0.18	0.25**	0.22*	0.15	0.09	0.42***	-0.05	—	
10.RECURRENT WORDS	0.75***	0.27**	0.2*	0.01	0.1	0.24**	0.1	0.22*	0.09	—
11.APPELLATIVES	-0.08	0.14	0.14	0.01	0.04	0.06	-0.08	0.11	0.12	-0.03

\* p < .05, \*\* p < .01, \*\*\* p < .001

Along with the correlation between repetition and elaboration, interpretation is highly correlated with elaborating the verbal flow of the episodes. Emotional and cognitive attributions of the infant's state are correlated with the elaboration process and are a significant part of the interpretation. Adjectives are the most prominent verbal use and seem to be part of this process through defining emotional and cognitive states. Recurrent words are significantly correlated with repetition and are also part of elaboration. Interestingly, using end markers to finish the episodes is highly associated with the attribution of the cognitive state of the infant.



### 6.3. Fundamental frequency and intensity graphics of narrative and control episodes

We have conducted analyses on graphics of fundamental frequency and intensity of French narrative and control episodes as well as Italian narratives. Graphics include separate acoustic curves for fundamental frequency and intensity for each episode.

We expected to observe an overall arc-shaped acoustic trend with a parallel increase and decrease of the fundamental frequency and intensity for French and Italian narrative episodes, exclusively distinctive from the acoustic trend of French control episodes. This analysis is performed based on the main hypothesis regarding acoustic features, claiming that fundamental frequency and intensity increase after the beginning and decline towards the end of vocal interaction between mothers and infants.

**Table 31**

*Codings of narrative acoustic arc model*

<b>Contingency Tables</b>			
<b>Groups</b>	<b>Outcome</b>		<b>Total</b>
	<b>Not Fit</b>	<b>Fit</b>	
French Narratives	48	16	64
French Control	55	5	60
Italian Narratives	46	18	64

We have calculated the difference between proportions for French and Italian narratives and French narratives and the control group. Italian narratives fit the acoustic arc model with an estimated proportion of 28% (95% CI [0.18, 0.40]) and French narratives with an estimated proportion of 25% (95% CI [0.16, 0.36]). The fitting model appears equally likely in both groups ( $p = .20$ ).

**Table 32***Proportions of fitting the narrative model for each group*

	Proportion	95% CI		OR	95% CI	
French Narrative	0.25	0.16	0.36	1.59	0.78	3.27
Italian Narrative	0.28	0.18	0.40			
French Narrative	0.25	0.16	0.36	3.70	1.44	9.52
French Control	0.08	0.03	0.18			

French control episodes fit the acoustic arc model with a low estimated proportion of 8% (95% CI [0.18, 0.40]). The likelihood of fitting the acoustic arc model increases by 3.70 when the episode is coded under the label of narrative in French data ( $p < .001$ ).

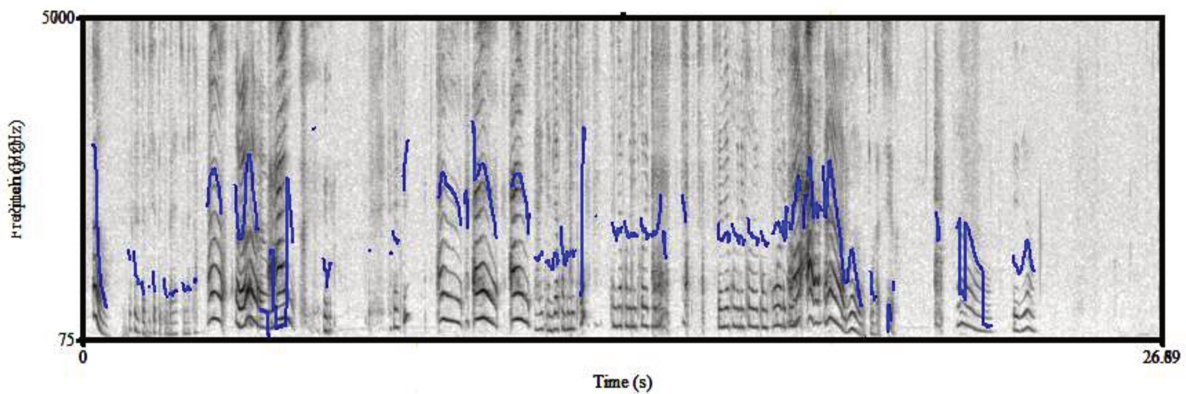
## 6.4. Case examples

This section provides examples from different narrative episodes along with acoustic images and verbal transcriptions. Spectrograms were obtained using Praat version 5.4.02 (Boersma & Weenink, 2014).

### 1. *French narrative in the recent past tense*

**Figure 17**

*Spectrogram of a French narrative in the recent past*



**Table 33**

*Transcription of a French narrative episode in the recent past*

Phrase	Transcription	Turns	Sequence
1	<i>Ah ! est-ce que ça t'a plu cette petite balade ce matin?</i>  Did you like the short trip this morning?	Mother	Introduction
2	<i>[vocalization]</i>	Infant	Introduction
3	<i>Ça t'a plu?</i>  Did you like it?	Mother	Orientation
4	<i>[vocalization]</i>	Infant	Orientation

5	<i>Oui tu aimes bien prendre le métro?</i>  Do you like to take the subway?	Mother	Development
6	<i>Tu aimes bien prendre le métro avec maman?</i>  Do you like to take the subway with mom?	Mother	Development
7	<i>Tu aimes bien prendre le métro dans le kangaroo avec maman?</i>  Do you like to take the subway in the kangaroo (baby carrier) with mom?	Mother (Overlapped vocalization following mother's speech)	Development
8	<i>[infant vocalization]</i>	Overlapped	Climax
9	<i>Ah bon! Tu me rassures</i>  Ah good! you reassure me	Mother	Resolution
10	<i>Ah! C'est bien</i>  Ah! This is good	Mother	Coda

The recording was made in the afternoon, the mother talks about a shared experience from the very recent past, and she asks her baby if he enjoyed the trip in the morning. The beginning is marked by a change of topic compared to the previous episode; the final pause is greater than 3000ms to display the ending of the episode. The use of simple present tense dominates the episode. The time indicator (Phrase 1: *ce matin*) demonstrates the time zone in which the event happened instead of the tense of the verbs. The mother interprets her baby's vocalization, which coincides with the climax, as an answer to her series of questions.

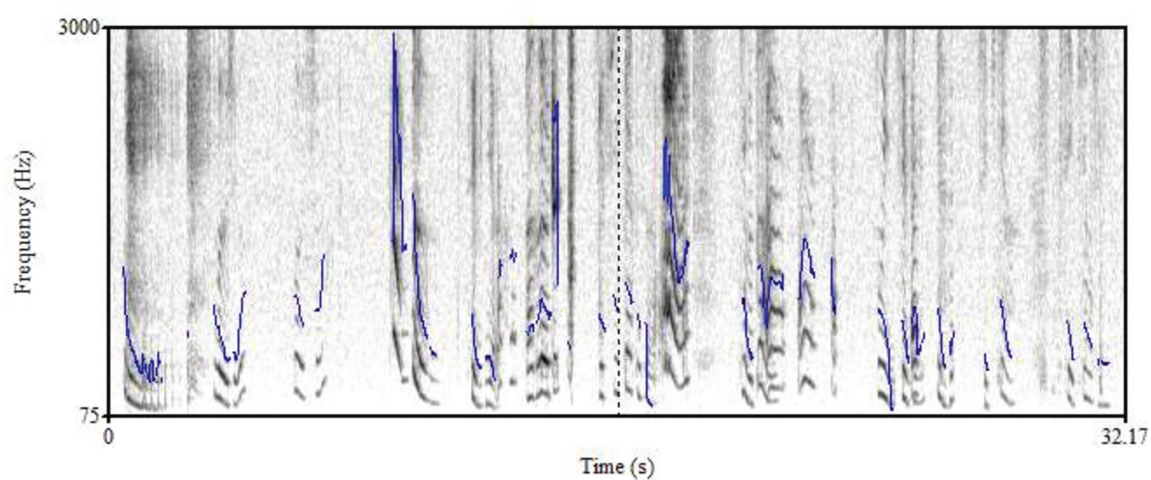
Acoustically, the recording starts with an ascending maternal vocalization. The maximum intensity value corresponds to the successive three infant vocalizations in bell shape. The output of maternal speech accelerates with repetition and variation of a question before reaching an overlapped climax marked by the infant's response.

Then, the rhythm of speech decreases slowly, and the episode comes to an end by the mother.

## 2. French narrative in the present tense

**Figure 18**

*Spectrogram of a French narrative in the present tense*



**Table 34**

*Transcription of a French narrative episode in the present tense*

Phrase	Transcription	Turns	Sequence
1	<i>(rire) Bonjour!</i> (laughter) Hello!	Mother	Introduction
2	<i>Bonjour!</i> Hello!	Mother	Introduction
3	<i>[vocalization]</i>	Infant	Orientation
4	<i>Ah!</i> Ah!	Mother	Orientation
5	<i>C'est bien!</i> That's good!	Mother	Orientation

6	[vocalization]	Infant	Development
7	Areuh! Areuh!	Mother	Development
8	[vocalization]	Infant	Development
9	C'est super tout ça! That's great!	Mother	Climax
10	[vocalization]	Infant	Resolution
11	Tu me parles! You speak to me!	Overlapped	Resolution
12	[infant vocalization]	Overlapped	Resolution
13	[vocalization]	Infant	Resolution
14	Oui, c'est bien Yes, this is good	Overlapped	Resolution
15	[infant vocalization]	Overlapped	Resolution
16	(rire) (laughter)	Mother	Coda

Here is a typical example of a present-tense narrative, based on an immediate interpretation of infant behavior as context. Mother is interested in the vocal expression of her baby, encouraging him with comments to vocalize more. Infant vocalization corresponding to the 8<sup>th</sup> phrase is a modest imitation of the mother's "areuh".

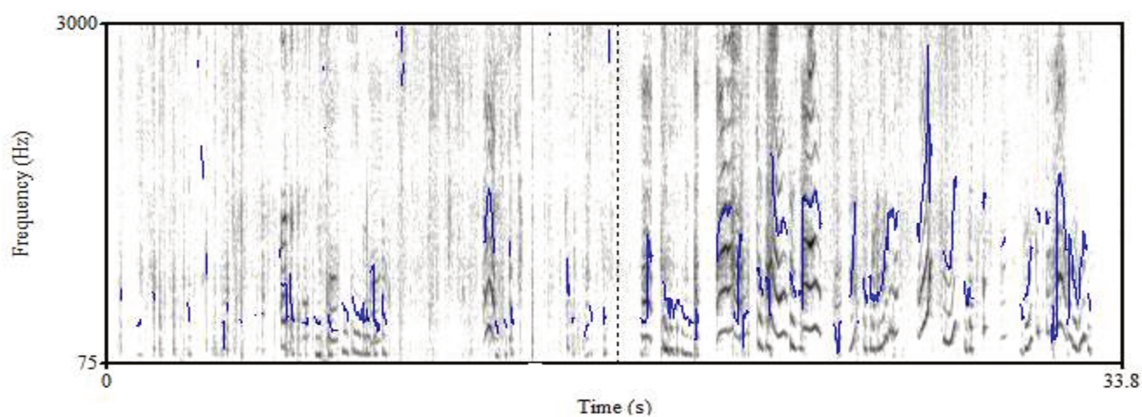
Regarding acoustic parameters, maximum intensity corresponds to the two successive infant vocalizations in the development part. There is a slight increase in the intensity of maternal vocalization, reaching to the climax, and it coincides with maximum frequency.

The episode corresponds to the beginning of the recording and is followed by another narrative episode with a shared topic.

### 3. French narrative in the future tense

**Figure 19**

*Spectrogram of a French narrative in the near future tense*



**Table 35**

*Transcription of a French narrative episode in the near future tense*

Phrase	Transcription	Turns	Sequence
1	<i>Attention!</i> Attention!	Mother	Introduction
2	<i>Hop hop ! Monsieur, Monsieur je vomis tout le temps.</i> Hop hop ! Mister, Mister, I vomit all the time.	Mother	Introduction
3	<i>Qu'est-ce que c'est que ça ?</i> What is this?	Mother	Orientation
4	<i>[infant vocalization]</i>	Overlapped	Orientation
5	<i>Qu'est-ce que c'est que ça?</i> What is this?	Mother	Orientation
6	<i>Quand est-ce que tu vas arrêter de vomir?</i>	Mother	Development

	When will you stop vomiting?		
<b>7</b>	<i>[vocalization]</i>	Infant	Development
<b>8</b>	<i>Quand est-ce que tu vas arrêter?</i> When will you stop?	Mother	Development
<b>9</b>	<i>[infant vocalization]</i>	Overlapped	Development
<b>10</b>	<i>Quand ça ?</i> When is it?	Overlapped	Development
<b>11</b>	<i>[vocalization]</i>	Infant	Development
<b>12</b>	<i>Quand ça ?</i> When is it ?	Mother	Development
<b>13</b>	<i>[infant vocalization]</i>	Overlapped	Development
<b>14</b>	<i>Quand tu boiras moins du lait.</i> When you drink less milk.	Mother	Development
<b>15</b>	<i>[vocalization]</i>	Infant	Climax
<b>16</b>	<i>Ah oui ! Tu bois beaucoup trop.</i> Ah yes! You drink too much	Mother	Resolution
<b>17</b>	<i>Beaucoup trop!</i> Too much!	Overlapped	Resolution
<b>18</b>	<i>[infant vocalization]</i>	Overlapped	Coda
<b>19</b>	<i>Oh oui!</i> Oh yes!	Overlapped	Coda

Mother comments on the current situation and its projected solution in the near future. We observe the use of near-future and future tense along with present tense because the event has double time-lapse; current behavior and future consequences of this behavior. We see in the discourse of the mother, examples of repetition and variation and also speaking on behalf of her baby. This can also be portrayed as an

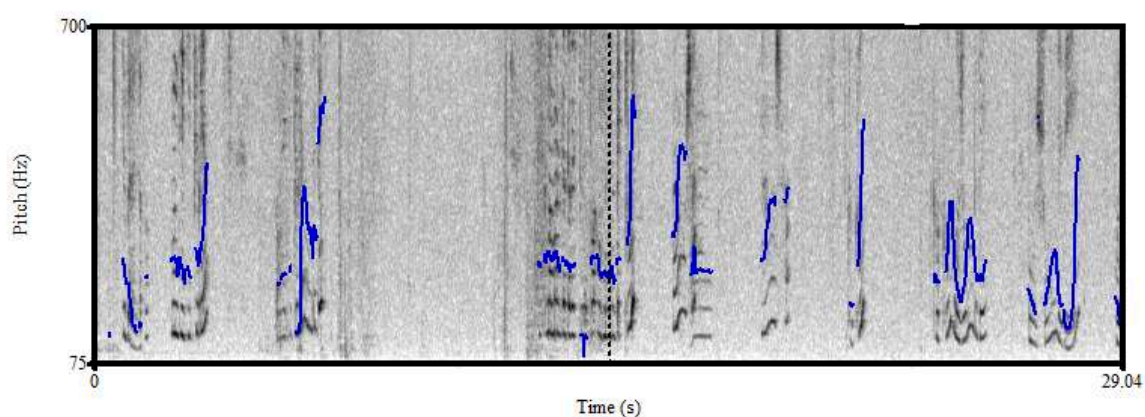


interpretation. Climax corresponds to an infant vocalization having one of the highest values in intensity. The output of the speech is accelerated in the development part compared to the beginning. This is the only narrative episode found in the recording.

#### 4. *French control episode*

**Figure 20**

*Spectrogram of a French control episode*



**Table 36**

*Transcription of a French control episode*

Phrase	Transcription	Turns	Sequence
1	<i>Eh bein alors?</i> What's up then?	Mother	
2	<i>On te dit plus rien ?</i> <i>No one tells you anything?</i>	Mother	
3	<i>On te dit plus rien mon bébé ?</i> <i>No one tells you anything my baby?</i>	Mother	

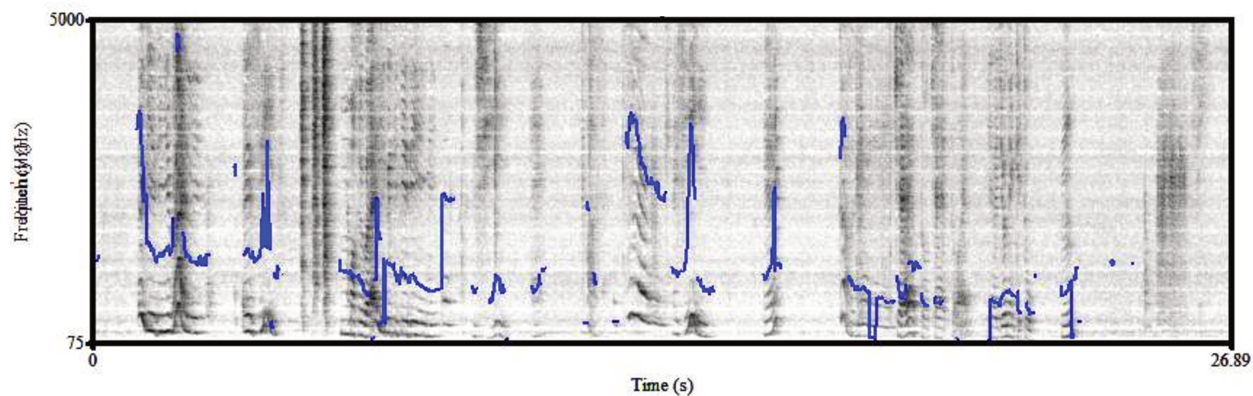
4	<i>Parce que maman ne s'intéresse plus à toi</i> <i>Because mom isn't interested in you</i>	Mother	
5	<i>Ah mon chat?</i> <i>Ah my kitty ?</i> <i>[vocalization]</i>	Overlapped	
6	<i>Ah, mon chat?</i> <i>Ah, my kitty?</i>	Mother	
7	<i>Coucou!</i> <i>Hi there!</i>	Mother	
8	<i>Coucou bonjour mon bébé !</i> <i>Hi there my baby!</i>	Mother	
9	<i>Bonjour mon bébé!</i> <i>Hello my baby!</i>	Mother	

In this case, we see examples of the use of repetition and questions like in narrative episodes. Baby hardly vocalizes within an overlapped vocalization. The baby is not as responsive and collaborative as the mother expects. Acoustically, the overall intensity contour traces a horizontal line with no major variation. As for the frequency values, it traces a descending overall contour and increases at the end. The duration of this episode corresponds to the narrative episode within the same recording. The reason why we call them non-narrative is to demonstrate the contrasts to evaluate the quality of the interaction.

## 5. Italian narrative episode

**Figure 21**

*Spectrogram of an Italian narrative episode*



**Table 37**

*Transcription of an Italian narrative episode*

Phrase	Transcription	Turns	Sequence
1	<i>Eh si ! Piedini belli</i> Oh yes! Beautiful feet	Mother	Introduction
2	<i>[vocalization]</i>	Infant	Introduction
3	<i>Eh si!</i> Oh yes!	Mother	Orientation
4	<i>Ma che belli questi piedini !</i> How beautiful are those little feet	Mother	Orientation
5	<i>[vocalization]</i>	Infant	Development
6	<i>Cosa? Eh si...eh cosa?</i> What is it? Oh yes...what is it?	Mother	Development
7	<i>[vocalization]</i>	Infant	Climax
8	<i>Eh si...eh cosa?</i>	Mother	Resolution

	Oh yes...what is it?		
9	<i>[vocalization]</i>	Infant	Resolution
10	<i>Certo certo eh... hai ragione anche tu...hmm</i>  Of course of course... you are right you too...hmm	Mother	Resolution
11	<i>[vocalization]</i>	Infant	Coda
12	<i>Eh si!</i>  Oh yes!	Mother	Coda

This episode is set in the present moment, and it is centered on the body parts of the baby. Mother comments on the body and interprets the vocal responses of the baby. The episode starts with the highest maternal discourse pitch values corresponding to high-intensity values. The climax, however, coincides with an infant vocalization and has a high score intensity value. The resolution includes an interpretation of the infant vocalization, and the ending is marked by lengthening the final syllable.

## Chapter 7: DISCUSSION

The present research aimed to identify and describe recurrent prelinguistic patterns in early mother-infant vocal interaction. Its main objective was to identify a co-constructed organization within the naturalistic vocal interaction between the mother and infant and to collect shared characteristics for the purpose of determining a taxonomy dedicated to identifying narrative organization. The research is framed by prior studies on narratives from many different disciplines. The narrative organization has long been a lens through which to comprehend textual organization. However, the study of oral storytelling opened a new standpoint while retaining core traits of text-based narratives: A three-level model including an introduction to the topic, development of this main topic with an apex point representing a turning point, followed by a brief conclusion.

### **7.1. Discussion of the results of the present study**

#### ***In search of a pattern***

This research aims to describe a vocal-based communication model in babies' early months following birth. Exploratory listening to hundreds of naturalistic audio recordings from 2 corpora of French and Italian dyads enabled the determination of preliminary criteria for selection of proto-narrative episodes. Episodes initially selected based on acoustic observation had to include non-fuss infant vocalization and to involve lasting mutual engagement between mothers and infants. A “narrative indicator” chart was devised on this basis and the vocal sequences from both collections of narrative episodes (EVOG and VERONA) were then re-coded based on it. A secondary collection of French vocal sequences served as a control modality to perform comparative acoustic and verbal content analysis.

Malloch (1999) presented an initial example of musical narrative in mother-infant vocal interaction by excerpting a familiar nursery rhyme that the mother sings to her baby during a vocal exchange. Malloch described narratives as the combination of pulse and quality units that unite rhythm and melody. In this example, he emphasizes

that the baby acts like a musical being by joining in consistently with her mother's singing and by introducing vocal variations to her mother's interpretation, marking her unique "musical style." The mother recognizes these variations and incites responses to her baby's style. This moment is a distinct example of how the infant perceives the original structure of a musical piece and how she can bring her own vocal signature, which makes this interpretation unique (Malloch, 1999:47).

This dialog presents a four-level organization, with the climax being the most significant phase of the entire interaction. The climactic moment of the interaction is described as the smallest meaning-making unit (Delafield-Butt and Trevarthen, 2015), which is claimed to be an essential moment contributing to the primary sketches of sensing the self. The interaction's ultimate meaningful moment includes both participants' emotions, energy, and intention. The climax appears due to subtle negotiations and delicate attunement of multiple forms of communication within the dialog.

Condon (1976) describes the nature of human communication as "fundamentally organized" (p. 285), and body movements and behaviors are shaped in sync with the conversation. Narrative forms such as songs, plays, and daily routines are already pre-structured in sequences. In this present study, we have exclusively focused on free vocal interaction because language is also an organizer. Language and vocalizations are the soundtracks of the embodiment of vocal expression. Still, the improvisational nature of free interaction challenges pinpointing a steady structure within a dynamic vocal stream. Therefore, it is more appropriate to consider early dyadic vocal interaction as music-like rather than language-like due to its formation by collaboration, just like two musicians' interaction during an improvisational session (Gratier, 2008). This complex process includes many different layers of communication, such as listening, adjusting, and responding, and gives rise to emergent cycles of joint engagement by association.

### ***French and Italian narrative episodes***

The French-Italian comparison was made, first of all, to validate the narrative model in two similar cultures. We have conducted this study to test the existence of

common patterns in each culture and to identify core features common to them. The main difference from previous mother-infant interactional studies is that we have considered not only local contingencies between both partners but also the organisation of discourse and exchange over time. Narrative episodes are created simultaneously over time, with the contribution of both parties. Thus, an episode should be seen as a whole rather than fragmented. The build-up of phrases and alternation and synchrony make these episodes unique compared to other periods of everyday social interaction.

Italian intra-group vocal dynamics comparisons displayed an overall trend to similar to French narratives. A significant difference between pause measures was detected in the comparison of both groups, indicating an age effect caused by the 8-to-12-week age interval in the French sample. A moderate variance was observed concerning measures of infant volubility: French infants scored high in almost all volubility indicators compared to Italian infants. Maternal vocal contribution measures suggest that Italian mothers spent more time vocalizing than French mothers, offering an explanation for the difference in infant vocal participation between the two groups.

#### ***Principal indicators of a narrative organization in early interaction***

We first devised a chart of indicators and coded all our pre-identified sequences accordingly. We then conducted a principal component analysis on the coded indicators from the French and Italian proto-narrative collections. Indicators were grouped under four labels, including the entire episode's global dynamic and the beginning, middle, and ending sequences. Similar results were observed for both groups, with small variations on certain items.

Moreover, the principal component analysis was performed to illustrate major indicators of the narrative indicator chart. As seen in Figure 6, results show that acoustic markers of fundamental frequency and intensity appear to be the principal indicators of narrative organization and are highly correlated between the four labels. In addition, repeated melodic contour, repeated group of words, and slow speech rate at the ending are further principal items indicating a narrative organization.

### ***Comparison of French narrative and non-narrative episodes***

Control protoconversational vocal sequences from the French sample allowed us to better study some of the internal patterns that characterize narrative vocal interaction between the mother and the infant. Our results show that internal patterns of narrative and control episodes differ. Infant volubility is one of the important measures confirming their difference. Regarding coordination and continuity, the number and duration of pauses also highlights this effect.

Narrative and control episodes were selected within the same audio recordings. In some cases, we identified that one narrative episode follows another, and the distinction was made based on a slight topic change. These episodes can be called latched narratives, and in terms of theme, they are related to the previous one. The same case also happens for control episodes, which are preceded by an episode labelled as narrative; however, in these episodes the coordination does not develop in such a way as to unite the two partners in a shared narrative objective.

Another important aspect of this comparison is whether there is an effect on vocal contribution, especially in terms of infant volubility (Iyer et al., 2016). This study aimed to explore mother-infant vocal interaction beyond infant-directed speech. We wanted to identify if a certain degree of organization in interaction motivated more infant vocal contributions. Our results show that both intra and inter-group comparisons of narrative and non-narrative episodes display different vocal dynamics. A strong positive relationship between infant and maternal vocalizations was detected within narrative episodes. While maternal vocal contributions do not differ in control episodes, infant volubility diminishes. Narrative episodes especially show a rise in vocal rates with a strong effect size. Overlapped vocalizations also display significant differences in comparison in terms of duration and number of occurrences. The number of pauses does not differ in comparison; however, they are more variable in terms of duration, indicating longer pauses within control episodes, which can be interpreted as an indicator of disengagement. These results support the idea that a particular organization in interaction, a narrative frame, leads to more contribution from the infant.



These results can also be interpreted in light of the early work of Murray and Trevarthen (1985). Their study showed the importance of infants' identification of contingency during the interaction. Our results show that maintaining contingency is not related to the amount of maternal talk but to the infant's interpretation. The feeling of connectedness becomes mutual when the infant apprehends the interaction as contingent and willingly participates. On the other hand, we must also recognize the fluidity of the interaction, knowing that social contingency is not the ultimate outcome of every interaction, especially in the early periods of life. Monitoring both narrative and control episodes assigns the amount of coordination variability within the experience and process of interaction. Some of the control episodes in our database are placed before or after a narrative episode and contain similar words or groups of words. This observational finding is an example of how social coordination is achieved through alternations of repeated interaction attempts. Both communication partners and the interaction probably underwent many scaffolding processes to sustain the interaction at different levels (De Jaegher & Di Paolo, 2008).

### ***Narrative analysis of the verbal content***

Verbal content demands specific attention to its meaningfulness. Here we are at the crossroads of verbal and nonverbal since we discuss meaning. When it comes to musical proto-narratives in mother-infant interaction, meaning is blended by both linguistic and musical semantics. Since the episode is co-constructed and one of the interactants has no access to the language, the meaning we extract also has an interactional dimension. It is a collaborative meaning-making process with a recognizable thematic texture based on language. We may think of proto-narratives as a means to the creation of meaning.

The verbal content analysis we conducted on this corpus is unrelated to the content described in IDS studies. Our content analysis is based on the narrative indicators conveyed in text and oral analysis. We tried to work out how mothers formalize verbal expression within a narrative episode. Repetitive words or groups of words appear as a common pattern in narratives, just like in poetry and music, but also

in IDS. Our primary purpose for this analysis is to show how the content shapes this specific form of interaction.

Verbal content analysis was conducted on four major categories determined by narrative indicators. The context of the episodes was examined in light of verbal tense use and time and place indicators. As shown in Figure 16, the present tense appears to be the dominant tense and is equally used in both narrative and control episodes. No tense use indicating utterances with no action words is only observed in control episodes. Place and time indicators are slightly represented in both groups, with time indicators making a moderate contribution in narrative episodes. Words indicating time and place have a greater place in defining narratives, especially in text analysis; however, in early speech, they appear to be weak indicators.

Theme and chronology categories contain elements of both narrative and infant-directed speech indicators. The use of questions and repetition is already claimed to be a significant quality of infant-directed speech by many researchers (e.g., Fernald et al., 1989). Therefore, they take place equally in both narrative and control episodes. However, an additional examination of repetitive words shows that non-narrative episodes could be an attempt of an introduction to a narrative episode, because some of them are immediately followed by a narrative episode. Some non-narrative episodes are placed close to the narrative episodes occurring within the same recording. They also have the same word or group of words matching the previous or following narrative episode.

Like time indicators, temporal order does not come forth as a strong indicator. The use of adjectives, elaboration of the speech, and interpretation of infant behavior arise as strong indicators in narrative episodes and are correlated with each other. End markers to conclude episodes also display significant features in narrative episodes compared to control episodes.

In the agency category, notable differences are spotted in the use of emotional and cognitive attributions, as well as appellatives, along with the use of you pronoun designated to the infant. These outcomes suggest that the infant is the main focus of

the conservation in narrative episodes and emits noticeable vocalized or nonverbal signals interpreted by the mother.

Traditional literature describes narratives with a beginning/middle/end model. Following this model, we wanted to observe differences from the classical narrative structure depending on the context and situation of the early interaction. What we mean by differences are the deviations from the classical model based mostly on text analysis. The main difference is the dominance of the present tense in narrative episodes. Classical narratives are known for the expansive use of past tense. However, in mother-infant interaction, we encounter mostly that the storyline is realized here and now. Stern (2004) says that the present moment is experienced as a lived story. It is a significant intersubjective experience told and lived in sync. It is a fragment in time plotted by “vitality affects” (Stern, 2010). Stern also describes the importance of the present moment through therapeutic relationships. “Rewriting” past experiences in the present time creates a change effect in the patient, and he changes his own story in real-time, giving it a new meaning and transforming the past meaning of the experience. He also mentions that the present moment corresponds to a phrase unit in dance, poetry, and between a mother and her baby. Experiencing these “moments of meeting” (2004, p.182) is what creates the narrative effect. For example, in French narratives, we come across the use of past tense, representing recent past experiences or near future plans between a mother and her baby. How could the present tense be considered a narrative? Huber (2016) shares an analysis of the use of present-tense narration by contemporary Anglophone writers. She reports that in the last century, the present tense has surpassed its “non-narrative and non-fictional” reputation and become “a tense of narration proper” (p.13). She emphasizes that simultaneity and present tense are connected; somehow, its usage was restricted for a long while to other genres like poetry rather than story. The use of present tense is wide in the sense of narrating past, present, or future. Considering the distinction between diegesis and mimesis, Huber says that the present tense creates its own plot in here and now in a way used in discourse rather than a story told by a narrator (p.23). In a similar term, Van Dijk (1975) distinguishes artificial narratives, narratives we encounter in stories or films, from natural narratives; generated spontaneously in conversation.

In “Tense and Narrativity” (1990), Fleischman says that using the present tense in text aims to neutralize collective norms and approach narrative discourse. Narratives in the present tense are more descriptive, based on the current situation and ambiance concerning people who are present. Past tense is related to the memory of a diegetic kind of storytelling; however, the present is the tense of performance.

Rather than recalling memories, the present is based on observation and reinterpretation, which slightly lapses between observation and speech time. Dahl (1985, cited in Fleischman, 1990) states that present-time narratives aren’t assigned to a specific time in the past; somehow, we can encounter some time indicators that are “minimal.”

### ***Acoustic Measures***

As fundamental frequency and intensity of vocal interaction are found to be important components of defining a narrative organization, this fits with our expectations on the musical qualities of this interaction. Acoustic analysis indicates that the overall frequency range does not differ when comparing the French narrative and control episodes. However, individual vocalizations show a significant narrative effect, indicating an increase of fundamental frequency in both maternal and infant vocalizations in French narrative episodes. The inter-language comparison also displays differences in the fundamental frequency, indicating both French mothers and infants show higher levels of fundamental frequency. This effect also appears in the overall comparison of fundamental frequency between French and Italian narrative episodes.

Regarding intensity, overall values are higher in narrative episodes than control episodes. Infant vocal intensity is higher and more variable in French narrative episodes than in control episodes. However, the overall intensity range of episodes shows that French dyads have a higher intensity range than Italian dyads.

The triadic acoustic comparison, including Italian and French narrative episodes and French control episodes, points out that some control episodes acoustically correspond to the narrative-arc model. This analysis shows that the pattern expected for narrative episodes appears in five recordings out of sixty control episodes. This

finding coincides with the observation indicating that along the continuum of the whole recording, some patterns appear to be seen in both narrative and control episodes, same as similar topics and words appearing within the same audio recordings. In other words, the distinction between episodes could be clearer-cut, which signals that the build-up of the narrative pattern is in a proto-period. Still, results show that the narrative episodes are more likely to fit the acoustic narrative-arc model even though the fit group represents a small proportion.

The acoustic narrative-arc model that we refer to, representing the simultaneous rise and fall of acoustic parameters from the beginning to the end of vocal sequences, is described as one of the essential markers of narrative vocal organization. However, this pattern was mostly studied through multimodal coordination during the interaction, considering vocal and behavioral outcomes from both partners (e.g., Delafield-Butt & Trevarthen, 2015). Our hypothesis was built on this description, even though our data was only composed of audio recordings. Besides, the inter-coder reliability of this study is acceptable; however, the alpha value is lower than the commonly accepted threshold, presumably due to the heterogeneous vocal dynamics within the episodes. The real-life data of acoustic measures relatively represent a rise and fall contour covering the whole episode, and internal vocal fluctuations make it difficult to interpret these acoustic graphics.

## **7.2. Discussion of narrative organization through the lens of early vocal interaction**

### ***The importance of studying proto-narratives in early social interaction***

The nature of expression, whether written or spoken, changes the internal dynamics of the story. As an example, Labov and Waletzky (1967) and Labov (2006) report that oral narratives begin with an opener serving as a synopsis at the beginning of the conversation. Many factors, such as the emotional intensity of the lived event and the audience, influence the storyteller in spontaneous conversations. Besides, the narrative organization is not only studied by linguists. Many neuroscientific and psychology studies suggest that our brain has the capacity to organize what we experience in a narrative way (e.g., Bruner, 1990). These studies go along with

developmental research, especially in how infants and children express themselves and extract meaning from the outside world. Language learning, expression, and interaction in early periods of life were studied to understand how they gradually become interactional partners. However, the preverbal period needed a different approach to examining this peculiar exchange where the infant has no access to full linguistic expression yet. The theory of Communicative Musicality (Malloch, 1999; Malloch & Trevarthen, 2009) suggested an innovative way to describe early interaction at the intersection of language and music. Authors define early vocal interaction as a period where musical and linguistic expressions are blended. From then on, the musicality of the vocal expressions turned into the core subject of prelinguistic studies. Thus, this research is centered on music-like features of early interaction and verbal content. Nevertheless, our study is equally engaged in how these vocal expressions are organized during spontaneous interactions. Prior research defined qualitatively how these interactions have been framed around a theme and the internal dynamics. We would like to conduct a quantitative research design to study these organization formats within the early mother-infant interaction. We have adopted a naturalistic approach to observe how the interaction spontaneously takes form. We also focused on an early period of life, the preverbal period, owing to less study dedicated to this critical stage to understand the premises of communication.

Another intention of this study is to monitor infants' vocal contribution and how it varies while specific organizational patterns are determined during social interaction. The qualities of infant vocal expression are crucial for future linguistic and social engagement and cooperation during the developmental course. Hence, special attention was brought to analyzing infants' participation in light of various indicators.

The study of natural vocal interaction brings to the fore many methodological questions bearing on the design of naturalistic studies. The taxonomy of narrative indicators was predetermined based on previous studies. The methodological challenge of defining a narrative format in vocal interaction was resolved by analyzing paired control episodes: infant participation became a focal point for comparing paired narrative and control episodes in the French corpus. The comparative study between

French and Italian dyads was also conducted to outline specific features exclusive to the narrative format.

Why is it important to study an early but essential period of time for the infant? We wanted to understand how vocalizations are bound to the social and cultural context within which they appear. Cultural-specific patterning and maternal styles were reported by numerous authors, and “motherese” does not appear universally in all cultures (Ochs & Schieffelin, 1984). Yet, to our knowledge, despite the methodological differences in studies, vocal variations of maternal discourse are described as musical in most early interaction studies.

On the other hand, prelinguistic vocal development has primarily been investigated in the light of developing linguistic skills. The majority of research has been carried out on the transition from the babbling stage to first words, for it is considered the first stage of language acquisition, and babbling lends itself to segmentation. Furthermore, it gives us strong clues on language pathologies. However, the infant begins to communicate long before his first syllables or words. We want to see how infants participate in the early meaning-making process through their mothers before linguistic form and meaning appear.

Along the way, we tried to operationalize descriptive analyses, question different degrees of narratives, and expand the research on oral narratives between verbal and nonverbal participants. These early intersubjective narratives represent an earlier format of an infant’s supported expression.

From the beginning to the end, melodic contours follow each other to construct a whole, a complete and unitary entity. Topic, frequency of pauses, maternal speech flow, and infant’s vocal contribution synchronize for creating a story in the present moment.

### ***Question of narrative thinking***

Julio Cortazar, in his book “Rayuela” (*Hopscotch* in English, *Marelle* in French), offers the readers two ways of reading: one is a traditional linear reading from the beginning to the end, and the other one is following the order that the author suggests

at the beginning of the book, which is jumping from one section to another. Either way provides different narrative effects. The second experience represents a fragmented narrative, a literary experiment rather than simple reading. Cortazar forces the reader to narrativize, think, and empathize with the protagonist's story. The book's name comes from the experience Cortazar proposes to his readers. This experience suggests that narrative forms are not always present or on display. And, we must try to extract and shape the meaning and create the narrative format.

The text rhythm also has its distinctive signature for writing procedures and the reader. Hustvedt (2018) reports that Virginia Woolf creates her phrases aloud while walking (p. 24). She emphasizes the distinction and the commonness of rhythm and time; while time is unfelt, we get vibes of the rhythm. The rhythm within the text, music, human actions, or even biological functions gives us some feeling that merges with a hint of familiarity, belonging, and being in life.

Considering past theories about narratives, we define musical proto-narratives according to our perspective. Musical "proto" narratives comprise three significant features: timing, verbal content, and musicality. In addition to these three features, musical proto-narratives are co-constructed by the mother and the infant in an early period and are still in progress. Musicians, especially in jazz improvisations, start by framing time or a theme in the progression of the musical piece. Musicians explore the sound like infants explore their voices. Then, they distribute the notes equally or not within this frame. Choosing a target note, a particular one that the musician wants to play during jazz improvisation, and playing passing notes to escalate some tension and reach the target note creates a narrative impact on the listeners. If we consider the target note as the climax of the solo, in the musician's mind, the musical phrase is first constructed from the climactic moment. Köhl mentions (2003) "the focal point" of a musical phrase, which might be the highest note, and demonstrates the pivotal point of a musical phrase. Similar to the example of interaction between jazz musicians in improvisation, "jazz mom" and her baby compose their own story-song. They are "playing good time." The recordings presented in this study are just snapshots of prolific interactions.



The organization of musical narratives is highly complex in analysis due to the intersection between its features. The distinction between musical discourse and the musical story is sometimes difficult to make. The dramatic development and progression are changing unexpectedly. In some cases, we encounter single or multiple plots. These variations take us to our initial questioning of the different degrees of narratives. These observations lead us to share a similar perspective defined by De Jaegher and Di Paolo (2007), positioning individual and social components of the interaction on a spectrum.

De Jaegher & Di Paolo (2008, 2007) suggest this approach to studying social interaction by focusing on autonomy, coordination, and sense-making while presenting a new concept called “participatory sense-making.” This view also emphasizes the importance of observing the self and the environment from a dynamic interactive point of view to understand social cognition beyond the nature-nurture dichotomy fully. The authors describe sense-making as an activity in constant interaction with the environment and with the self. It is a perpetual cognitive process that continually regulates self-using mechanisms like autonomy and coordination. The organism is not a passive agent but an active being that experiences and makes sense of it. From this point of view, coordination must also be considered a dynamic and changing process. We see many different forms of coordination in social interaction situations (e.g., turn-taking, overlapped vocalizations) or musical exchange (e.g., repetition or imitation of a musical phrase). So, the coordination could not be assumed as total synchrony because regulation plays a big part in this process. As coordination evolves, specific parameters between interactional partners become structures, leading to this social interaction's autonomy. Autonomy in social interaction represents the separate selves and the interaction itself. Here, we see that the autonomy of social interaction also changes participants' autonomy within the frame of the specific social interaction. On the other hand, the autonomy of individual participants in the interaction becomes involved in the sense-making process through coordination. This is what authors define as a “participatory sense-making” process.

De Jaegher & Di Paolo (2007) emphasize the fluidity of social interactions, saying that coordination is a variable activity. The narrative pattern, especially in the early period, is dynamic. It appears in different degrees, with a greater range of expression.

### **Timing**

Timing is one of the primary elements of narrative structure appearing under the label of sequences. We considered the timeframe of the narrative episode, the temporal crossings within the episode, the temporal structure of the whole episode, and musical time as narrativizing is put in an order in time. Stern (2000) insists on repeating some aspects of a mother's speech, such as words, syllables, or gestures, to maintain her baby's attention and suggests that these repetitions are highly regular, in a word rhythmic. Repetition creates some sort of temporal mastery. In addition to the repetition topic, we also come across some examples of signature patterns in the mother's speech. As for imitation, when one creates a pattern, the other picks it up with slight modifications, either by repetition or variation.

Repetition gives the music a sense of coherence and maintains unanimity of the musical meaning. It stresses musically and meaningfully significant motifs, like in the example of the chorus of a song. Innate consistency of repetition enables anticipation and allows variation afterward.

Besides, variations derived from these repetitions forge the speech even more musical. Consecutive sequences of repetitions and variations frame the temporal experience. Stern (2000) states that the mother's speech variation is intentional. The infant's vocalizations also vary but are not as clear as the mother's. The baby makes variations over variations, and their speech grows musically.

The rhythm of their speech and the alternation of inner rhythms create a baseline. These are the primary sketches of adjustment forms of regularity. Musical phrases and melodic contours can be seen as primary chunks by cutting the time intervals out to organize the memory narratively. The static and the dynamic, the constant, and the variable, are interchanging along the creation of interaction.

Like the no-gap-no-overlap principle in turn-taking, temporal coordination is crucial in social interaction to maintain coordination and make sense. Temporal coordination gives a structural baseline to the interaction, allowing the partners to make necessary adjustments and regulations to preserve the dynamics of making sense (De Jaegher & Di Paolo, 2007).

Although a steady rhythm creates a sense of harmony, some vivid moments in interaction can create derivations from regularity to express more powerful emotions. In a paper examining the dynamics of rhythmic applause of spectators in concert halls, Néda and colleagues (2000) found that spectators reduce their clapping frequency to achieve a common rhythmic synchronization. Otherwise, as the intensity of applause increases, the synchrony gets lost; however, it might be adjusted again by decreasing the overall frequency distribution of clapping. Thus, the rhythmic baseline is no longer synchronized when the intensity escalates, and vice versa. This back-and-forth of applause harmony is shaped by the spectators' agreement on how they want to demonstrate their appreciation. Therefore, the degree of enthusiasm induces variations in synchronicity, resulting in another form of expression emphasizing intense admiration. The latter is similarly observed in the case of narrative episodes between the mother and the infant, where the interaction reaches a climactic point. The acceleration of different acoustic measures, such as frequency, amplitude, and speech rate, describes this sequence.

### **Musicality**

The third and most significant feature of this particular interaction is musicality. Similar to the verbal content, the musicality of the episodes is collaborative as well. In this case, the prosody of the mother's speech and the musically meaningful vocalizations of the infant are in prospect, both separately and conjointly. There is a kind of heterophony in one melodic gesture. Along the way, we witness how mother and baby adjust to each other musically. It progresses through pitch organization on a horizontal melodic line. Rhythmic movements shape the skeleton and the frontiers of this musical interaction. To examine the musicality within this interaction, we first identified the climax because it is "the unit of meaning-making (Delafield-Butt &

Trevarthen, 2015). Then, we discern initiation, growth, and decline parts in the vivid present at the crossroads of past, present, and future. For infants, this is a discovery process of self and others through vocal interaction. Forming narratives aims to reach the other and then share the moment.

Musical phrases scaffolding narrative episodes are micro-narratives. A musical phrase is a brief musical idea and gradually becomes a complete musical manifesto, with its melodic scaffold and rhythmic baseline along the harmonic background. Narratives are phraseological, containing sequential repetition and variation.

Schoenberg (1967), in his book “Fundamentals of Musical Composition,” makes the analogy of music without organization to an essay without punctuation (p.1). Some phrases only have ascending lines, but that doesn’t make them less narrative. When we look at the whole musical composition, we can find later the descending line that completes it. The same assessment is valid for Schoenberg’s works. He only presents his statements abandoning accustomed compositional systems. Schoenberg says, “Musical thinking is subject to the same dialectic as all other thinking” (p.94).

### ***Musical meaning and meaning-making process***

We can contemplate narratives in mother-infant interaction as a medley of musical and linguistic signs, providing a baseline for the future. Narratives have a semiotic function that is drawn away from verbal semiotics. It’s a semiotic process signifying intersubjective engagement.

Narratives are co-constructed by alternating or overlapped vocalizations of the mother’s and baby’s intersubjective sharing experience. From a micro perspective, separate melodic contours trace an inflection of both mothers' and infants' interacting voices within the narrative episodes. These melodic contours of each vocalization get unified and lead to an arc-shaped musical gesture – **an audible gesture** (Trevarthen, 2000). In other words, these vocal fluctuations turn into a common gesture organized in time. This gesture arises from a precocious musical bonding between adults and infants. These patterns represented in mind build up primary sketches of meaning, proto-semiosis.

## ***Re-Thinking Musical Narratives***

“Music emerges from silence; music is interrupted  
or temporally disturbed silence.”

Jankelevitch (1961:170, cited in Tarasti, 1994).

We alternately call musical proto-narratives in mother-infant interaction “floating narratives,” inspired by the term “floating intentionality” of Ian Cross (1999). This implies the meaning-making process drifted between the two communication partners, meanwhile weaving them together. Questioning the narrative structure at this period of a lifetime, we suggest that it is a pattern in progress alongside its improvisational nature and its cultural, musical, and social motives that represent the narrative socialization of the infant.

In musical composition performance, let’s say that the oeuvre reflects the past or future dreams of the compositor in the present time. Composed in the past, performed in the present, and sometimes modified in the present or the future, the composition is influenced by the present moment actors or happenings. The dialogical function of the present tense blends with the past composition and creates a new narrative. The meaning that the compositor attributes to his work and the meaning that listeners afford to the work are blended in this new one. As we see in the example of John Cage’s 4’33”, altering the surroundings of the present moment introduces a constantly shifting character to the composition and again becomes a new narrative anyway.

Tarasti (1994) says minimalist music is anti-narrative (p.285), defined as the transgression of conventional narrative forms. He says, “Minimalist music rejects temporal segmentation and traditional functions of musical time: one can no longer distinguish among beginning, ending, introduction [...]” (p.284).

Unlike the structure of traditional classical music, minimalist music is marked by its continuity and repetition as a striking unit in its composition. In the classical view, the score is created here and now, even if every narrative is related to past living. The present’s presence affects every musical piece’s evolutionary step, whether performed,

written, or imagined. Musical narrativity is still the subject of great debate but opens new perspectives to think about narratives differently from the classical structure.

On the question of the plot without language in music, mother-infant interaction makes a mid-way between language and music, providing plot, but mostly vocally. The idea of narrativity appears as a meta-structure. We consider that the first aim of making music is not telling a story but “saying something,” which becomes a story. The narrativity is a latent component of a musical piece, and the narrative in music arises as chunks to the listeners appearing in the whole score or during the improvisation. Tarasti (1994) suggests that French semioticians would say narrativity should be investigated on the level of musical utterances and performance (p.23).

The pitch reflects the variation within the flow and reaches a climactic point. However, intensity also creates a path similar to pitch variations. Usually, in correlation, intensity, and frequency trace different paths from time to time. Lee & Schögler describe three steps of the evolution of intensity: attack, sustain, and decay (2009, p.96). In the attack phase, intensity increases immediately and falls gradually, then rapidly. The tension brings a movement to progress; the energy created from this tension is released. The level of energy and movement depends on the composer while writing, or in the case of performance, it depends on the musicians' interaction or even the audience's participation.

The basic narrative structure is introduction/development/conclusion, derived from literature, a humanly-created work of art. The literature also contains different forms of expressions like poetry, which contains a different internal dynamic than a story; however, it is also a narrative.

In the 20<sup>th</sup> century, serialism in music laid the first stone on a new musical path. It has a brand new effect on the listener. However, serialism or twelve-tone music has its own peculiar pattern for the composition and the listener. The creating process evolves by choosing a series of notes. These notes may change in octaves to observe pitch jumps that create significant variation. These are the techniques often used by serialist musicians while composing; however, no doubt serialists like to break the rules. The question of sounding right is neglected in serialism; the music genre's importance is the

composer's self-expression and will to liberate the strict composition. From the listener's perspective, they meet the unexpected while listening to Schoenberg's work. This might create a sense of instability and uncanny feelings. Breaking the expected patterns opens new perspectives to those dealing with vicious circles of what was learned. This effect is similar to the release effect of narratives. Atonal music has a different narrativization process, forcing us to think differently.

### **7.3. Limits of the study**

From the beginning of this research, we have encountered some limits and difficulties along the way. First of all, we conducted a limited number of statistical analyses due to inter-individual differences in the occurrence of narratives. Even when comparing narratives and control episodes, despite the duration equivalence of the episodes, the number of vocalizations differing within and between episodes made the analysis procedure complicated. An overall acoustic comparison is only performed based on graphs representing acoustic measures.

In this thesis, we studied a period corresponding to the beginning of discernible infant vocalizations. This choice also comes with its difficulties. Much research in the field of mother-infant interaction focuses on the babbling stage when transcription is more straightforward and vocal units are clearly discernible, so less research has been done on the early vocal production of the infant in interaction with the mother. And so forth, we insisted on working on this early period despite its requirement of great effort since this early stage is fruitful in terms of musicality and communicative signs.

The recordings containing more than one task for mothers in our initial corpus of French dyads limited the duration of free interactional moments and possibly the emergence of narrative episodes. The variable recordings conditions suggest that the more the dyad is relaxed and free from tasks, the more they have quality time. In the Italian sample, each dyad had a ten-minute free interactional moment. Thus, the recording conditions were different between the two corpora and may have induced some important biases.

#### 7.4. Perspectives for future research

In terms of future research perspectives, the most important one should be considering the interaction's multimodality. Future studies should focus on all communicative signs produced by infants and will be complementary and integral to interpreting narratives, such as gaze, smile, and arm movements. These nonverbal behaviors usually appear in maternal discourse while she interprets her baby's reactions. However, the relationship and coordination between the bodies will certainly shed light on the narrative nature of the interaction as a whole. Overlapped vocalizations are particularly interesting in terms of synchrony; however, they are hard to study in terms of acoustic analysis. Recordings made in an isolated environment can be useful to observe the dynamics of these simultaneous vocal units but they may lose out on ecological factors.

Secondly, a longitudinal study should follow the evolution of narrative patterns at least during the first six months. It is essential to observe how narratives take shape and evolve how durations and vocal units differ, and how the transition from early vocal to the babbling stage is made. Such research certainly needs teamwork and collaboration. It is also important to consider physiological measures connected to infant responses throughout an interactional period. This could provide a biological basis for the questions that we ask. A combined methodology, including neuroimaging methods and advanced sound analysis tools, may be helpful to examine different aspects of early human interaction fully and to identify recurrent patterns of communication rigorously. Additionally, a further study including naive listeners to evaluate narrative features of interaction will allow the assessment of how they perceive the qualities of the interaction.

Validation of the tool is also one of the important aspects of applying it to a larger population. This will allow us to conduct this analysis among different cultures. In 2014, I conducted a pilot study on Cameroonian mother-infant dyads filmed in a rural area<sup>5</sup>. Video recordings weren't made specifically for acoustic analysis. So, I just conducted a global analysis based on listening. This study helped me to observe how cultural-

---

<sup>5</sup> The author would like to thank Professor Heidi Keller for giving her access to her data collection.



specific properties and music influence our interactional skills. African music is best known for its rhythmic structures. These mother-infant dyads that I have listened to created some repetitive patterns while interacting, and *groove* appears at the climax of their conversation. It was more music-like and nonverbal compared to the European dyads' interaction. Additionally, body contact and parenting styles differ from the European context. That is to say, despite the audio quality, which doesn't allow us to work on those recordings properly, this exploratory observation helped me immensely to see how culture and music are profoundly linked.

An additional study on the narrative arc hypothesis might also help to understand the acoustic dynamic. An extended version of our analysis on fundamental frequency and intensity graphs coded by naive listeners will shed light on acoustic parameters and their variability throughout narrative and non-narrative episodes.

The Theory of "Communicative Musicality" offers a theoretical background to investigate interaction musically and gives us a foreground to inspire future research. All the components of this present research still need some more investigation.

Does sound itself express something? Sound is how first encounters are made with the outside world. It starts in utero, and the meeting becomes multimodal with multiple people after birth. Sound and musicality are universally used by people who want to reach each other. My work with children with autism opened new ways of communicating. This experience favoured thinking deeply about communication disorders, as well as early periods of interaction. As sound is the first touch with the world, we become interested in following the sound to see what it bears. We have presented a minimal frame of a pattern provided by the interaction and extended from the sound.

## CHAPTER 8: CONCLUSION

This thesis contributes to the pattern identification process by recognizing individual and social components of a dynamic human interaction system. The approach covers the study of individual and interactive vocal components by exploring the aspects built around indicators similar to verbal narrative organizations. A naturalistic action plan was adopted to examine everyday social interaction in infants' home environments.

Narrative and story are usually accepted as synonymous, representing an account of events. However, a story predominantly indicates a series of connected events with a strong connotation to written forms. Nowadays, narratives include an extensive way of organization and a variety of expressions. The verb storytelling, on the other hand, incorporates a wider perspective of expressions as well. In tandem, narratives are presented in various forms of expression, such as written, oral, and nonverbal. Their analysis provides post hoc structural characteristics. Such characteristics display organizational markers like temporality and agency in connection to an additional aspect of narratives, the experience.

The human experience of narratives has been subject to various interdisciplinary research encompassing cognitive processes and verbal and nonverbal expressions. Oral narratives or storytelling are described as performances by many authors, including the audience. Interactive narratives such as music or proto-narratives in early interaction are the most recent forms of interest in narrative studies. The latter is essential for its improvised and spontaneous nature and for using multimodal expressions. Narratives have their own diachronicity, like language; hence, the early interaction provides prominent aspects to the origins of narratives.

The portrayal of narratives presents various aspects. One aspect that is common for all narrative researchers is the markers of a beginning and an ending. Many authors describe different levels or degrees of narratives and present various ranges. Narratives without a plot set an example of a minimal narrative form. In this case, the

plot seems to be the subject of discussion; however, it is considered the most salient aspect of the narrative structure since the beginning of narrative studies.

The perspective from which we observe storytelling, i.e., focusing on the storyteller or considering storytelling as a performance, profoundly influences our definition within the narrative range. Verbal expressions naturally afford an episodic structure to the story. The aesthetic features of the verbal flow add dimension to narrative expression and experience that we observe in the cases of oral narratives and interactive narrative forms. Together with nonverbal expressions, they enhance emotional articulation and prompt a narrative performance.

The musicality of vocal expressions was in focus to emphasize the peculiar nature of mother-infant interaction, defined as music-like by many researchers before the time. We emphasize musicality in mother-infant interaction because their vocal exchange is co-constructed similarly to how musicians interact with each other, which differs from conventional linguistic conversations. The present research centers on how individual vocalizations by both partners are co-constructed to form a story-like organization. The identification process shows that early vocal interaction is more than finding synchrony; multiple attempts of coordination and association are in place. Therefore, we suggest that, as we see in all interactional frames, this interpersonal engagement is an active process rather than a well-defined structure. Yet, vocal expressions gather through some organizational qualities. This organization, which we call proto-narrative, has value to be studied by researchers to comprehend the grounds of communication through the lens of nonlinguistic signs and early meaning-making moves. Thus, these proto-narratives certainly have more emotional qualities than comparable interaction episodes.

A quantitative analysis of musical proto-narratives in early mother-infant vocal interaction still needs extended research. Proto-narratives in early vocal interaction should be considered a process incorporating narrative-like structures and represent primary grounds for future storytelling frames and formats. We expect this thesis to proliferate future methodologies to scrutinize this challenging material.

## **De la narrativité verbale à la narrativité non verbale des interactions mères-bébés :**

### **Proto-narratifs musicaux dans l'interaction précoce**

#### **Résumé**

Inspirées par Jérôme Bruner (1990) qui préconise la prise en compte de l'activité narrative comme une capacité précoce de l'esprit humain organisant les expériences sous formes de narrations, nous portons notre attention sur les premiers échanges parlés entre mères et nourrissons parmi des populations européennes. Le bébé communique expressivement avant ses premières paroles. Les études minutieuses sur le bébé de moins de 6 mois en interaction nous montrent que le bébé est capable de communiquer, notamment avec sa mère, dès les premières semaines de vie (Malloch, 1999; Gratier, 2001). Avant l'acquisition du langage, il s'agit d'une communication précoce à travers une musicalité partagée et interactionnelle.

## Prologue

Alors que les approches les plus classiques de la narrativité ont été orientées sur les textes et les discours (Propp, 1928; Labov, 1967), les anthropologues et les psychologues soulignent la place importante que semble avoir l'activité narrative dans la vie humaine (Bruner, 1990; Ochs, 2002). Malloch et Trevarthen (2009) nous présentent une approche moins flagrante de ce concept dans la musicalité communicative des interactions entre la mère et son bébé, issue du développement psychologique du nourrisson et l'ont définie comme un apprentissage de l'expérience du temps à travers les échanges vocaux. L'interaction vocale entre la mère et son bébé avant la période du babillage fait l'objet de plusieurs recherches en psychologie du développement. Ces études nous permettent d'approfondir nos connaissances sur le développement précoce du langage et les prémisses de la communication chez le bébé. Les échanges vocaux ludiques deviennent de plus en plus élaborés et contribuent à l'instauration d'un réel dialogue entre la mère et le bébé même si, au début de la vie, ce dernier n'est pas encore un « virtuose » de ses cordes vocales. Cet article vise à présenter l'idée qu'une narrativité largement non verbale, au croisement du discours maternel et des vocalisations émises par la mère et le bébé, constitue une dimension centrale des interactions vocales précoces entre la mère et son bébé de 2 à 3 mois au cours de situations de la vie quotidienne.

Nous nous intéressons particulièrement à la musicalité des échanges vocaux et, dans ce contexte, la narrativité offre un potentiel unifiant entre le verbal et le non verbal, le contenu et l'expression affective co-construite. Dans le présent article, nous interrogerons la notion générale de narrativité sachant qu'elle apparaît sous différentes formes et avec diverses organisations, selon la formulation du contenu et du contexte (écrit ou oral). En d'autres termes, il est possible d'identifier qualitativement différentes structures narratives d'après un ensemble d'indices spécifiques. L'étude de la narrativité textuelle et orale nous amène à approfondir et à retracer le potentiel de cette notion dans le cadre des interactions spontanées entre parents et bébés et à développer l'idée d'une narrativité non verbale à partir de l'analyse des données provenant de dyades mères-bébés françaises et italiennes.

## **Le contexte des interactions précoces**

A l'âge de 2 à 3 mois, le développement cognitif et affectif du bébé s'achemine en parallèle avec le développement des compétences sociales. Comme par exemple, dès le premier mois, les bébés indiquent une préférence pour la voix de leur mère et pour leur langue maternelle et ils sont sensibles aux changements prosodiques de leur discours (Boysson-Bardies, 1996). De nombreux travaux ont été menés, surtout dans les contextes européens et nord-américains, sur les compétences communicatives du nourrisson et sur sa participation active dans des échanges vocaux et expressifs dès le deuxième mois de la vie (Gratier, 2001; Beebe et al., 1985; Papousek & Papousek 1989; Trevarthen 1977). Ces auteurs rapportent qu'à partir de cet âge, les bébés émettent des vocalisations claires (que l'on peut distinguer des sons végétatifs) et mélodiques. Ces échanges vocaux, appelés aussi « protoconversations », sont composés de vocalisations produites en alternance ou en simultanéité par l'adulte et le bébé au cours d'une relation de face à face visuel (Fogel, Toda & Kawai, 1988; Stern et al., 1975). Selon ces études, l'adulte modifie spontanément son style de parole lorsqu'il s'adresse au bébé. Ce langage adressé au bébé a été l'objet de très nombreuses études. La comparaison inter-langues montre par exemple d'importantes similarités entre les contours intonatifs et le discours de mères européennes, nord-américaines et japonaises lors d'échanges avec leurs bébés (Fernald et al., 1989). Quelques recherches sur des populations dans d'autres contextes culturels soutiennent l'existence de modes de communication vocale similaires mais présentant aussi des variations, comme par exemple en milieu rural en Inde du Nord et du Sud (Gratier, 2003; Keller, 2007; Stork, 1986), en milieu urbain au Nigeria (Trevarthen, 1988) ou en Afrique du Sud (Cowley et al., 2004). À titre d'exemple, lors de l'interaction avec leurs bébés, davantage de chevauchements ont été observés chez les mères indiennes. Ainsi, le temps de pause entre les tours de paroles est plus court chez les mères indiennes que chez les mères françaises (Gratier, 2003). Le langage adressé au bébé mais aussi l'interaction vocale qu'il stimule semblent ainsi présenter à la fois des caractéristiques communes dans diverses communautés linguistiques et culturelles et des styles spécifiques qui ont pu être reliés aux pratiques et représentations ayant cours dans ces contextes culturels (Gratier, 2003; Fernald & Morikawa, 1993). Les caractéristiques principales du langage

adressé au bébé sont un tempo ralenti, une plus grande articulation des unités sonores, des contours mélodiques répétés et une variation importante de la fréquence fondamentale de la voix (Fernald & Simon, 1984). A la lumière de ces travaux sur l'interaction préverbale chez le bébé, la théorie de la musicalité communicative apporte un nouvel éclairage pour révéler les strates de ces échanges en posant l'hypothèse que ceux-ci peuvent être analysés comme des phénomènes musicaux. Ainsi, la musicalité communicative se lit par la modélisation musicale de l'interaction vocale entre le nourrisson et ses parents (Malloch, 1999; Malloch & Trevarthen, 2009). La notion de musicalité communicative prend en compte toutes les variations vocales et corporelles du bébé qui expriment des émotions et du sens lors des contacts avec ses partenaires. Cette théorie nous permet d'analyser ces échanges selon trois paramètres : 1) la pulsation, les intervalles entre les unités de temps et successions régulières de ces unités ; 2) la qualité, représentée par les contours mélodiques et les mouvements corporels ; 3) la narrativité, la combinaison des unités de pulsation et de qualité qui s'imbriquent pour façonner des séquences expressives et significatives. Ces trois dimensions se rejoignent et apparaissent sous forme d'un dialogue musical manifestant une expressivité à la fois individuelle et collective, tout en s'inscrivant dans une temporalité spécifique et propre aux capacités communicatives du bébé d'âge précoce.

### **La narrativité au croisement du langage et de la musique**

Les séquences d'une narrativité verbale, décrites de manière similaire chez plusieurs auteurs (Propp 1928; Todorov 1969; Labov, 1967) incluent: 1) une introduction, donc l'initiation au sujet, aux personnages, puis 2) un développement dans lequel sont présentées des actions orientées vers un problème ou un moment de crise, et enfin 3) une conclusion, c'est-à-dire la résolution du problème ou de la source de tension. En s'appuyant sur ce modèle, il est possible de décrire la structure narrative provenant d'un texte, d'un récit, d'une histoire racontée ou d'un film. De plus, un récit expose clairement les éléments et les facteurs mettant en évidence son déroulement ordonné qui projette une fin, ses événements inattendus, situant des personnages dans des contextes. Le sens du récit dépend en grande partie de sa dimension verbale. Pourtant à l'oral, les histoires racontées spontanément manifestent des structures

désordonnées et atténuées (Labov, 1967). Dans ses travaux sur la narrativité des conversations de la vie quotidienne, Ochs (2002) souligne le fait que ces narrations influencent nos façons de penser et d'interagir et sont elles-mêmes influencées par les systèmes culturels de connaissances, de croyances et de valeurs. Ochs met l'accent sur la dimension collective et sociale des narrations orales, reconnaissant l'importance de l'aspect interactif entre les interlocuteurs pour composer ces narrations. Comme Bruner (1990), elle suggère que les narrations sont essentielles pour concevoir les contextes culturels du développement humain et qu'elles soutiennent le processus de socialisation de l'enfant dès le plus jeune âge. Ochs affirme qu'il existe différents niveaux de narrativité. Ainsi, le contenu verbal des récits spontanés ne serait pas le seul moyen de raconter les événements : les mouvements du visage et du corps, les gestes et les contours prosodiques contribueraient autant à l'instauration d'une mise en tension narrative. On y retrouve peut-être les fondements de toute une narrativité expressive dissociée du langage, comme dans la danse, le mime ou la musique. De plus, dans les formes narratives non verbales, sans l'apport du langage, les éléments structurels ne sont pas toujours évidents. Comment donc conceptualiser la structure narrative d'une pièce musicale ? La répétition de motifs pourrait constituer un élément important de la narrativité musicale. En effet, les répétitions sont utilisées fréquemment dans le contexte musical et sont moins présentes dans les textes littéraires, sauf si celles-ci sont des expressions poétiques. Selon Imberty (2005), l'œuvre musicale s'organise autour de répétitions et de variations, et ces fragments, organisés ensemble, permettent une structuration du temps. En d'autres termes, la succession des différents segments d'un morceau musical édifie «la continuité d'une forme entière» (Imberty, 2007: 8) dans le flux temporel. Cette dynamique de continuité-discontinuité à l'intérieur des répétitions et des variations n'est peut-être pas le propre de la musique elle-même, mais pourrait spécifier une narrativité implicite qui peut se manifester dès le début de la vie humaine. Dans l'interaction vocale entre la mère et le bébé, chaque énoncé se décompose en motifs. Ces motifs sont liés entre eux à plusieurs niveaux. Il s'agit d'une interaction constante et multiple entre diverses dimensions de la production vocale: rythme, hauteur, intensité. Le discours maternel adressé au bébé est lui-même porteur d'une narrativité verbale et d'une narrativité non verbale. En effet, dans le langage adressé au bébé, l'adulte fait un usage plus



important des indices prosodiques et des formes sonores onomatopéiques. L'analyse de cette narrativité adressée au bébé a d'ailleurs révélé que le bébé est lui-même acteur et participant de son organisation temporelle, composée de répétitions et de variations, de pauses et de vocalisations simultanées, qui identifie ce que l'on peut appeler un « épisode narratif » dans l'interaction parent-bébé. En effet, dès les premières semaines de vie, le bébé vocalise à des moments particuliers du déroulé des énoncés maternels et la mère en retour ajuste son expression à l'intervention du bébé (Gratier & Trevarthen, 2008). Cette dimension poly-interactive se révèle aussi dans l'alternance des tours de paroles des partenaires. L'ordre et l'alternance des unités vocales, y compris les pauses qui interviennent, s'imbriquent et tracent le contour musical qui contient un sens co-construit. L'organisation séquentielle des expressions réciproques de la mère et du bébé fait émerger une trame temporelle qui confère à l'échange une qualité affective particulière. La narrativité donne une ampleur proto-sémiotique à toute production vocale du bébé et maintient la cohésion du discours de la mère. Le sens préverbal prend naissance dans la musicalité des échanges. Cela peut s'observer d'ailleurs lors d'une improvisation entre musiciens de jazz par exemple (Gratier, 2008): une exploration sonore mutuelle se déploie progressivement dans le temps et façonne une création d'ensemble. Tenant compte de nos connaissances sur la narrativité, nous avons procédé à des analyses acoustiques d'enregistrements de dyades mère-bébés afin de pouvoir décrire le plus finement possible l'organisation narrative dans l'interaction spontanée au croisement du verbal et du non verbal.

### **Collection des séquences**

Des séquences d'interaction libre entre des mères et des bébés (16 bébés français âgés de 8 à 12 semaines et 16 bébés italiens âgés de 12 semaines) ont été enregistrées à domicile. Toutes les mères ont pour langue maternelle le français ou l'italien selon l'échantillon observé. Les familles des deux échantillons ont un niveau socio-économique de classe moyenne intermédiaire ou supérieure. Environ 240 minutes d'interaction des dyades françaises et 200 minutes des dyades italiennes ont été analysées. Les extraits choisis comportent des séquences dans lesquelles la mère parle en continu à son bébé, les pauses entre vocalisations ne dépassant pas 3 secondes (Van Egeren et al., 2001). Trois secondes de pause sont acceptées comme un marqueur

de contingence sociale dans une interaction dont le bébé est un participant actif. Dans les séquences sélectionnées le bébé vocalise au moins une fois. Le début et la fin des séquences ont été identifiés soit par des pauses supérieures à 3 secondes, soit par un changement de sujet dans le discours maternel. La cohérence sémantique permet de délimiter le contenu verbal afin de déceler les indices narratifs non verbaux des séquences. La méthode d'analyse acoustique de ces interactions a été adaptée à partir de celle décrite par Malloch (1999) et Gratier (2001, 2010). Elle permet d'obtenir un ensemble de mesures pertinentes sur les dimensions musicales (par exemple la fréquence fondamentale, la hauteur, le tempo, la durée...) du langage adressé au bébé, à l'aide de logiciels d'analyses spécialisés. Nous avons utilisé le logiciel Praat pour obtenir les mesures acoustiques (cf. Chapitre Observation Acoustiques). Chaque unité vocale a été codée séparément avec un codage spécifique pour le discours de la mère, la vocalisation du bébé, les pauses entre les vocalisations des deux participants et les vocalisations simultanées. Notre analyse se centre sur des moments spécifiques dans le déroulement des interactions vocales entre la mère et son bébé. Nous supposons que ces derniers se réunissent autour d'un thème précis, marqué par des séquences à partir desquelles nous nous attendons à observer une organisation narrative (introduction-développement-conclusion) au niveau vocal. Le thème du discours maternel lors de l'interaction est composé des répétitions d'un son, d'un mot et/ou d'un groupe de mots selon les réponses du bébé.

Cette organisation permet de décomposer des épisodes dans lesquels nous constatons différents types de formes vocales de la voix maternelle et des vocalisations des bébés. Nous avons donc analysé les échanges des dyades françaises et italiennes afin de pouvoir distinguer les similarités et spécificités des styles narratifs de l'interaction vocale dans les deux cultures. Nous avons choisi deux langues maternelles d'origine latine dont les composantes syntaxiques et grammaticales sont similaires.

## Description des observations recueillies

### Observations acoustiques

Pour décrire la dynamique partielle et globale des séquences narratives, nous avons déterminé, à partir des extraits sélectionnés, des indicateurs pour chaque séquence en fonction d'analyses qualitatives réalisées avec le logiciel Praat version 5.4.02 (Boersma & Weenink, 2014). La dynamique partielle se décompose ainsi : – L'introduction comporte des contours mélodiques ascendants et est délimitée par une pause courte avant le début de l'épisode ou un changement de sujet par rapport à l'épisode précédent. – Le développement englobe l'alternance rapide de vocalisations, l'accélération du débit de la parole et/ou des vocalisations simultanées. En termes d'acoustique, apparaissent la hauteur maximale de la fréquence fondamentale et l'augmentation de l'intensité. – Pour la partie conclusion, nous avons remarqué des indices de clôture : la chute de la fréquence fondamentale, des contours mélodiques descendants, le ralentissement du débit, la chute de l'intensité et l'allongement des syllabes.

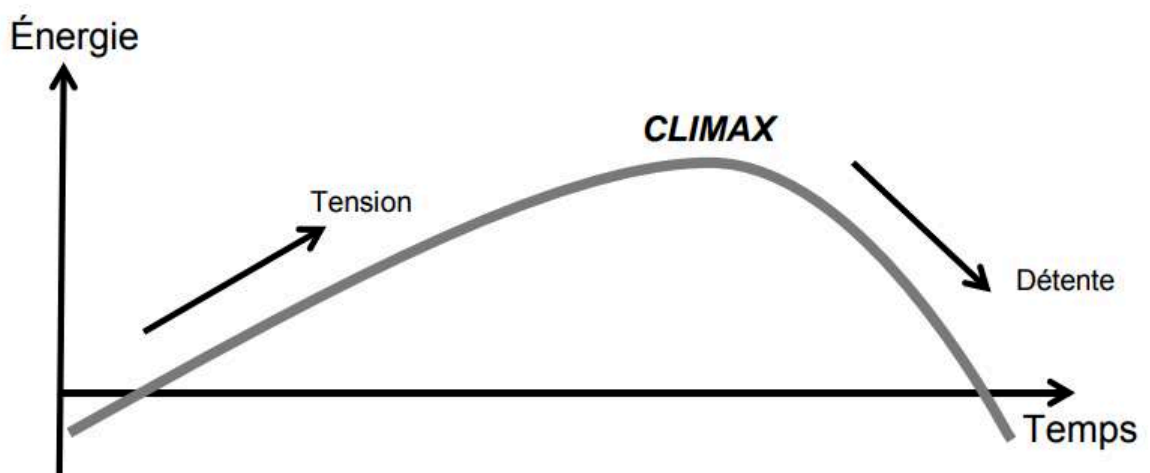


Fig. 1. Évolution de la dynamique narrative en fonction de l'énergie et du temps.

Quant à la dynamique globale, pour l'aspect acoustique, nous observons une courbe générale de la fréquence fondamentale et de l'intensité en cloche et en ce qui concerne le contenu verbal, la répétition d'un thème, d'un mot, d'un groupe de mots ou d'un son. Pour résumer, l'épisode narratif débute par une introduction, l'engagement mutuel de l'attention et une incitation; ensuite elle est suivie d'une séquence de développement où il y a un engagement émotionnel, une ligne de tension (Labov, 1972) marquée par des vocalisations simultanées et/ou alternées (climax) et enfin l'épisode se termine par un moment de ralentissement, la conclusion. Du point de vue acoustique, les séquences narratives se façonnent avec des grandes variations de fréquence fondamentale et d'intensité et ces fluctuations concrétisent musicalement l'originalité de ces séquences. Cela crée un contour spécifique sous forme d'arc mélodique (fig. 1).

### **Observation du contenu verbal**

Le contenu verbal est concret et méta-communicatif. Le discours de la mère se construit à partir des réactions spontanées du bébé et prend forme avec ces dernières. L'énoncé ne peut être interprétable qu'en tenant compte des paramètres non verbaux de l'interaction. Dans la plupart des situations observées dans les deux contextes culturels, les mères commentent les expériences actuelles, souvent en fonction de l'environnement immédiat, pour donner un sens aux mouvements et aux actes du bébé auxquels elles se synchronisent. Le thème du discours maternel est centré sur le corps et les expressions vocales, faciales et gestuelles du bébé. Bien que le moment présent soit dominant dans leur discours, les mères françaises ont aussi tendance à citer des expériences vécues du passé récent et du futur proche de leurs vies quotidiennes (tableau 1). La mère énonce un récit autobiographique des moments partagés avec son bébé même si le récit personnel de la mère englobe le récit personnel du bébé. Dans ce cas, le contenu verbal est relativement plus proche de la structure narrative d'un récit concernant le positionnement et l'évolution dans le temps. Dans tous les cas, le protagoniste du discours maternel est toujours le bébé. Nous reconnaissons les limites de cette approche basée sur l'analyse acoustique des échanges. Cependant, le langage non verbal possède plusieurs dimensions comme les mouvements corporels ou les mimiques. Ainsi, les comportements des bébés comme

les mouvements des bras ainsi que les signes non verbaux comme le sourire apparaissent dans le discours de la mère. En d'autres termes, la mère interprète et attribue le sens aux réactions du bébé.

*Exemple 1: L'épisode narratif d'une dyade française*

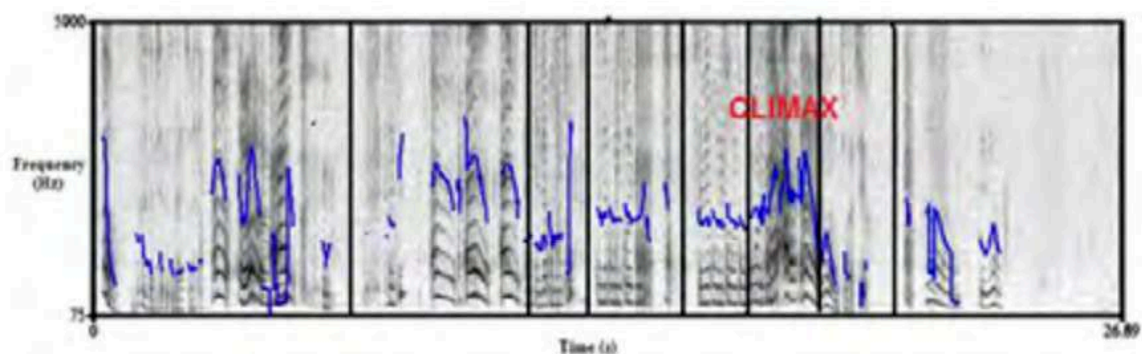


Fig. 2. Image sonore de l'épisode narratif d'une dyade française.

Énoncé	Transcription	Séquence
1	Ah est-ce que ça t'a plu cette petite balade ce matin ? [vocalisation du bébé]	Introduction
2	Ça t'a plu ? [vocalisation du bébé]	Introduction
3	Oui tu aimes bien prendre le métro ?	Développement
4	Tu aimes bien prendre le métro avec maman ?	Développement
5	Tu aimes bien prendre le métro dans le kangaroo avec maman ?	Développement
6	[vocalisation du bébé]	Climax
7	Ah bon tu me rassures	Résolution
8	Ah c'est bien	Conclusion

Tableau 1. Transcription de l'épisode présenté en fig. 2.

L'épisode débute avec un changement de sujet par rapport à l'extrait qui le précède et la fin de l'épisode est définie par une pause supérieure à 3 secondes. L'enregistrement a été fait dans l'après-midi. La mère parle d'un moment dans la matinée qu'elle a partagé avec son bébé (passé récent). Elle commence par une question générale sur la petite balade de la matinée, ensuite elle continue avec une question plus précise dans la partie développement. Elle répète plusieurs fois les mêmes mots, ce qui accentue le

thème et répète des groupes de mots avec une variation légère ou en accélérant le débit de parole. Cette séquence atteint le 'climax' par la vocalisation du bébé qui a été interprétée par la mère comme une réponse à sa question. L'épisode arrive à la fin par les commentaires de la mère aux réactions du bébé, la mère exprime sa satisfaction quant à la contribution vocale de son bébé.

Les mesures acoustiques sont présentées dans le tableau ci-dessous pour les valeurs moyennes et maximales de la hauteur de la voix et de l'intensité:

FR		Introduction	Développement	Conclusion
Hauteur (Hz)	Moyenne	183,67	356,43	229,70
	Maximum	250,45	432,44	270,94
Intensité (dB)	Moyenne	53,87	65,27	59,34
	Maximum	63,95	69,01	64,07

Tableau 2. Valeurs de la fréquence fondamentale et de l'intensité de l'épisode narratif de la dyade française.

*Exemple 2: L'épisode narratif d'une dyade italienne*

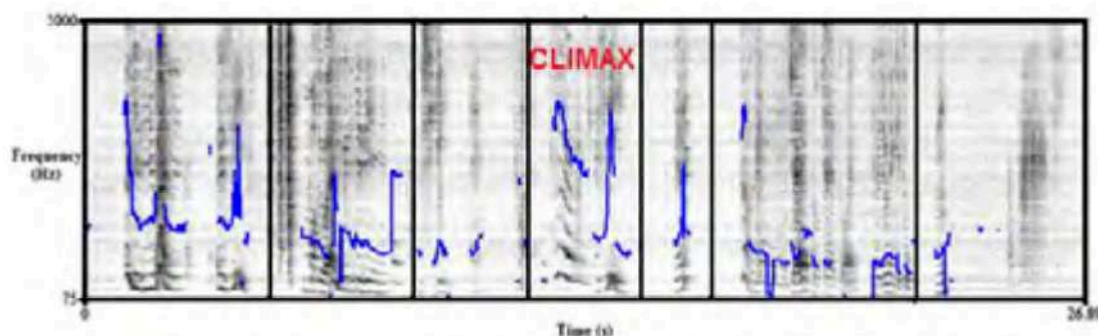


Fig. 3. Image sonore de l'épisode narratif d'une dyade italienne.

Énoncé	Transcription	Séquence
1	<i>Eh i piedini belli... Eh si</i> Eh les beaux petits pieds... eh oui	Introduction
2	<i>Ma che belli questi piedini</i> Mais comme ils sont beaux ces petits pieds [vocalisation du bébé]	Développement
3	<i>Eh si... eh cosa ?</i> Eh oui... eh quoi ?	Développement
4	[vocalisation du bébé]	Climax
5	<i>Eh si eh</i> Eh oui eh	Résolution
6	<i>Cosa ? Certo certo eh... hai ragione anche tu</i> Quoi ? Bien sûr bien sûr eh... tu as raison toi aussi	Résolution
7	<i>Eh si</i> Eh oui	Conclusion

Tableau 3. Transcription de l'épisode présenté en Fig. 3.

Tout au long de l'enregistrement, la mère commente les différentes parties du corps du bébé. L'épisode précédant cet extrait comporte des remarques sur d'autres parties du corps alors qu'ici, la mère se concentre sur les pieds. Dans son discours, la mère répète un groupe de mots avec une légère variation. Ensuite, elle encourage son bébé à vocaliser plus en lui adressant une question. Puis, le bébé répond avec une voix qui monte et qui correspond au climax de cet épisode. La mère indique avoir reconnu sa réponse. Les mesures acoustiques sont présentées dans le tableau ci-dessous :

IT		Introduction	Développement	Conclusion
Hauteur (Hz)	Moyenne	258,52	362,87	175,48
	Maximum	513,69	525,07	182,83
Intensité (dB)	Moyenne	67,55	59,57	59,75
	Maximum	70,22	66,90	61,61

Tableau 4. Valeurs de la fréquence fondamentale et de l'intensité de l'épisode narratif de la dyade italienne.

Le fait de choisir deux langues d'origine latine et similaires au niveau syntaxique et grammatical n'a pas pour but de comparer les deux cultures, mais de pouvoir observer les points communs de l'interaction précoce. Ces observations nous ont permis d'établir une topologie des critères que nous avons définis à partir des théories présentées en introduction de cet article. De fines variations ont été observées entre les deux contextes culturels ainsi qu'à l'intérieur des échantillons sur le plan interindividuel. La dynamique acoustique marque des différences légères, même parmi les différents enregistrements de la même dyade. Il s'agit de différents degrés de narrativité et il faut souligner que, dans le cadre d'une interaction précoce entre la mère et son bébé, la narrativité est plus un processus interactif qu'une structure ou un état stable et déterminé. Entre le 2<sup>ème</sup> et le 3<sup>ème</sup> mois de la vie, nous observons un ajustement au niveau vocal entre la mère et le bébé, que l'on peut qualifier de dialogue narratif. Derrière l'aspect verbal soutenu par le discours de la mère qui en manifeste le sens, l'aspect musical révèle aussi le sens d'un accordage affectif (Stern, 2003). Ces narrativités musicales semblent agir comme des régulateurs importants des états du bébé. Elles participent également à l'engagement mutuel et aux manifestations des pensées avant l'acquisition du langage. Ces narrativités musicales peuvent être aussi considérées comme les origines biologiques et culturelles de la socialisation précoce.

## **Épilogue**

L'interaction sociale préverbale entre le bébé et la mère a été décrite de manière à la fois musicale et narrative : les mères s'adressent à leur bébé d'une façon particulière sur le plan vocal avec des variations de hauteur et de rythme, et des fluctuations mélodiques. Lors de ces engagements mutuels, leurs discours se présentent sous différentes formes narratives : chansons, berceuses, comptines ou histoires fictives ou vécues. Reprenons la description de la narrativité selon nos observations : l'épisode narratif dans l'interaction vocale entre la mère et son bébé commence par une introduction. Cette introduction établit une conjonction attentionnelle qui se développe, à un moment de partage mutuel et émotionnel, par le chevauchement ou la succession rapide de vocalisations, et qui se termine par une diminution d'énergie. Les tours de parole s'emboîtent et forment par la suite un geste commun organisé



dans le temps. Ces proto-narrations sont les prémices d'une pensée emblématique<sup>6</sup> et cette expérience rudimentaire prépare le bébé à des expériences communicationnelles plus complexes a posteriori. Ceci n'est possible qu'avec la mise en séquence temporelle du sens. Ainsi, le geste musical (Kühl, 2011) illustre parfaitement le sens implicite d'un phrasé musical et les dynamiques incarnées dans le son. Pour Kühl, le geste musical est une représentation dans l'esprit humain qui se transforme par la suite en geste intériorisé. Notons que l'évolution de la construction du sens ainsi que son expression sont progressives. Cette dernière prend de l'ampleur dans le temps et avec le temps. Au croisement de plusieurs modalités expressives, l'interaction précoce est dotée d'une forme narrative particulière et constitue l'expression naturelle d'une dyade adulte-bébé.

---

<sup>6</sup> Représentation précoce de l'organisation de la pensée et du sens.



## REFERENCES

- Almén, B., (2008). *A Theory of Musical Narrative*. Bloomington, IN: Indiana University Press.
- Baker-Ward, L., Bauer, P. J., Fivush, R., Haden, C. A., Ornstein, P. A., & Reese, E. (2007). Coding coherence in autobiographical narratives. *Symposium conducted at the biennial meetings of the Society for Research in Memory and Cognition*, Lewiston, ME, USA.
- Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of real-world event schemas during narrative perception. *Journal of Neuroscience*, 38, 9689–9699.
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, 95, 709–721.
- Barthes, R. (1953/1972). *Le Degré zéro de l'écriture*. Paris: Seuil.
- Barthes, R. (1966). Introduction à l'analyse structurale des récits. In: *Communications*, 8, *Recherches sémiologiques : l'analyse structurale du récit*. pp. 1-27.
- Barrett, M.S. (2011). Musical Narratives: A Study of a young child's identity work in and through music-making. *Psychology of Music*, Vol 39, Issue 4, pp. 402-423.
- Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.
- Bateson, M.C. (1975). Mother-infant exchanges: the epigenesis of conversational interaction. *Ann N Y Acad Sci.*, 263: 101-13.
- Bateson, M. C. (1979). The epigenesis of conversational interaction: A personal account of research development. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 63-77). London: Cambridge University Press.
- Beebe, B., Jaffe, J., Feldstein, S., Mays, K., and Alson, D. (1985). "Interpersonal timing: the application of an adult dialogue model to mother-infant vocal and kinesic interactions," in *Social Perception in Infants*, eds T. Field and N. Fox (Norwood, NJ: Ablex Publishing Corporation), 217–247.

- Bergeson, T. R., & Trehub, S. E. (2007). Signature tunes in mothers' speech to infants. *Infant Behavior and Development*, 30, 648-654.
- Berliner, P.F. (1994). *Thinking in jazz: The infinite art of improvisation*. Chicago: The University of Chicago Press.
- Benveniste, E. (1939). Nature du signe linguistique. *Acta Linguistica*, 1(1), 23–29. doi:10.1080/03740463.1939.10410844
- Blacking, J. (1973). *How Musical is Man?* Seattle: University of Washington Press.
- Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, 14(2), 211-227.
- Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language*, 15(3), 469-480.
- Boersma, P. & Weenink, D. (2014). *Praat: doing phonetics by computer [Computer program]*. Version 5.4.02, consulté le 26 November 2014 URL: <http://www.praat.org/>
- Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. (2015). Mother-Infant Contingent Vocalizations in 11 Countries. *Psychological science*, 26(8), 1272–1284. <https://doi.org/10.1177/0956797615586796>
- Bornstein, M. H., Putnick, D. L., Suwalsky, J. T., Venuti, P., de Falco, S., de Galperín, C. Z., Gini, M., & Tichovolsky, M. H. (2012). Emotional Relationships in Mothers and Infants: Culture-Common and Community-Specific Characteristics of Dyads from Rural and Metropolitan Settings in Argentina, Italy, and the United States. *Journal of cross-cultural psychology*, 43(2), 171–197. <https://doi.org/10.1177/0022022110388563>
- Brandt, P.A. (2004b). *Spaces, Domains, and Meaning. Essays in Cognitive Semiotics*. European Semiotics Series, Volume 4. Bern: Peter Lang.
- Brandt, P.A. (2009). Music and how we became human – a view from cognitive semiotics: Exploring imaginative hypotheses. In S. Malloch & C. Trevarthen (Eds), *Communicative Musicality: Exploring the basis of human companionship* (pp 31-44). Oxford: Oxford University Press.

- Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum, *The effect of the infant on its caregiver*. Oxford, England: Wiley-Interscience.
- Brès, J. (1994). *La narrativité*. Duculot, Louvain-la-Neuve.
- Briggs, J.L. (1970). *Never in Anger: Portrait of an Eskimo Family*. Cambridge: Harvard University Press.
- Bruner, J. S. & Sherwood, V. (1976). Early rule structure: The case of 'peekaboo.' In J. S. Bruner, A. Jolly & K. Sylva (eds.), *Play: Its Role in evolution and development*. Harmondworth : Penguin.
- Bruner, J. & Feldman, C. (1993). Theories of mind and the problem of autism. In S. Baron Cohen, H. Tager-Flusberg, D. J. Cohen (Eds.), *Understanding other Minds: Perspectives from Autism*. Oxford University Press
- Bruner, J. (1985). The Role of Interaction Formats in Language Acquisition. In: Forgas, J.P. (eds) *Language and Social Situations*. Springer Series in Social Psychology. Springer, New York, NY. [https://doi.org/10.1007/978-1-4612-5074-6\\_2](https://doi.org/10.1007/978-1-4612-5074-6_2)
- Bruner, J. (1990). *Acts of Meaning*. Cambridge: Harvard University Press.
- Bruner, J. (1991). The Narrative Construction of Reality. In *Critical Inquiry*, 18(1), 1-21.
- Bruner, J., & Lucariello, J. (2006). Monologue as Narrative Recreation of the World. *Narratives from the Crib*. Ed. Katherine Nelson. Cambridge: Harvard University Press. 73–97.
- Brown P. (1998). Conversational structure and language acquisition: the role of repetition in Tzeltal. *J. Linguist. Anthropol.* 8, 197–221.
- Brown, H. D. (2000). *Principles of Language Learning and Teaching*. White Plains, NY: Longman.

- Carra, C., Lavelli, M., Keller, H., Kartner, J. (2013). Parenting Infants: Socialization Goals and Behaviors of Italian Mothers and Immigrant Mothers from West Africa. *Journal of Cross-Cultural Psychology*, Vol.44, Issue 8, pp. 1304-1320.
- Chatman, S. (1978). *Story and discourse, Narrative structure in fiction and film*. Ithaca, Cornell Univ. Press.
- Cheour-Luhtanen, M., Alho, K., Sainio, K., Rinne, T., Reinikainen, K. et al. (1996). The ontogenetic earliest discriminative response of the human brain. *Psychophysiology*, 33(4):478-81.
- Clark, H. H., & Brennan, S. A. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (eds), *Perspectives on socially shared cognition* (pp. 127-49). Washington: APA Books.
- Clark, E. V. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clarke, E. F. (2005). *Ways of listening: an ecological approach to the perception of musical meaning*. Oxford: Oxford University Press.
- Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving musical time. *Music Perception*, 7(3), 213–251. <https://doi.org/10.2307/40285462>
- Condon, W. S. & Sander, L. S. (1974). Neonate movement is synchronized with adult speech: Interactional participation and language acquisition. *Science*, 183, 99-101.
- Condon, W.S. (1976). An analysis of behavioral organization. *Sign Language Studies*, 13, 285-318.
- Cooper, R.P. & Aslin, R.N. (1989). The language environment of the young infant: Implications for early perceptual development. *Canadian Journal of Psychology*, 43(2), 247-265.
- Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61, 1584-1595.
- Corbeil, M., Trehub, S. E., & Peretz, I. (2013). Speech vs. singing: infants choose happier sounds. *Frontiers in psychology*, 4, 372. doi:10.3389/fpsyg.2013.00372.
- Corbeil, M.E., Trehub, S.E., & Peretz, I. (2015). Singing Delays the Onset of Infant Distress. *Infancy*, 21(3), 373 – 391.

- Cowley, S.J., Moodley, S. & Fiori-Cowley, A. (2004). Grounding signs of culture: Primary intersubjectivity and social semiosis. *Mind, Culture and Activity*, 11(2), 109-132.
- Cox, J.L., Holden, J.M. & Sagovsky, R. (1987). Detection of post-natal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, 150, 782-786.
- Cross, I. (2003). Music and biocultural evolution. In M. Clayton & T. Herbert & R. Middleton (eds), *The cultural study of music: a critical introduction* (pp. 19-30). London: Routledge.
- Cross, I. (1999). Is music the most important thing we ever did? Music, development and evolution. In S. W. Yi (Ed.), *Music, mind and science* (pp. 10-39). Seoul: Seoul National University Press.
- Damasio, A. (1999). *Le sentiment même de soi*. Paris: Odile Jacob.
- Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. Pantheon/Random House.
- DeCasper, A.J., & Fifer W.P. (1980). Of human bonding: Newborns prefer their mother's voices, *Science*, 208, 1174-1176.
- DeCasper, A. & Spence, M.J. (1986). Newborns prefer a familiar story over an unfamiliar one. *Infant Behavior and Development*, 9, 133-150.
- DeCasper, A., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, 17, 159-164.
- Delafeld-Butt, J.T. (2018). The Emotional and Embodied Nature of Human Understanding: Sharing narratives of meaning. In C., Trevarthen, J., Delafeld-Butt and A-W., Dunlop (eds.), *The Child's Curriculum: Working with the natural voices of children*. Oxford: Oxford University Press.
- Delafeld-Butt, J.T. & Trevarthen, C. (2015). The ontogenesis of narrative: From movements to meaning. *Frontiers of Psychology*, 6:1157 doi: 10.3389/fpsyg.2015.01157.

- De Jaegher, H., & Di Paolo, E. (2008). Making sense in participation. An enactive approach to social cognition. In Morganti F, Carassa A & Riva G (Eds.), *Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interactions*. Amsterdam: IOS Press.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507. <https://doi.org/10.1007/s11097-007-9076-9>
- De Léon, L. (1998). The emergent participant : Interactive patterns of socialization of Tzotzil (Mayan) children. *Journal of Linguistic Anthropology*, 8(2), 131-161.
- De Léon, L. (1999). Verbs in Tzotzil (Mayan) early syntactic development. *International Journal of Bilingualisme*, Volume 3, Issue 2-3, pp. 219-239.
- De Léon, L. (2011). Language socialization and multiparty participation frameworks. In A. Duranti, E. Ochs, & B. Schieffelin (Eds.), *The Language Socialization Handbook*. Blackwell Publishing.
- Demuth, C. (2008). *Talking to infants : how culture is instantiated in early mother-infant interactions. The case of Cameroonian farming Nso and North German middle-class families*. Unpublished Dissertation. University of Osnabrück.
- Dissanayake, E. (2000). *Art and intimacy: How the arts began*. Seattle, WA: University of Washington Press.
- Dunn, J. (1988). *The beginning of social understanding*. Harvard University Press.
- Duranti, A., & Burrell, K. (2004). Jazz improvisation: A search for hidden harmonies and a unique self, *Ricerche di psicologia*, 3, 71-101.
- Feldman, R. (2007). Parent–infant synchrony: Biological foundations and developmental outcomes. *Current directions in psychological science*, 16(6), 340. [doi.org/10.1111/j.1467-8721.2007.00532.x](https://doi.org/10.1111/j.1467-8721.2007.00532.x)



- Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental science*, 13(2), 271–278.  
<https://doi.org/10.1111/j.1467-7687.2009.00890.x>
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8(2), 181-195.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10(3), 279-293.
- Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American Mothers' Speech to Infants. *Child Development*, 64, 637-656.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborn. *Developmental Psychology*, 20, 104-113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477-501.
- Ferguson, C.A. (1964). Baby talk in six languages. *American Anthropologist* Vol. 66, No. 6, Part 2: *The Ethnography of Communication*, pp. 103-114.
- Ferguson, C.A. (1977). Baby talk as a simplified register. In Catherine E. Snow & Charles A. Ferguson (eds.), *Talking to Children*. Cambridge University Press. pp. 209-235.
- Fitch, W.T. (2010). *The Evolution of Language*. Massachusetts: Cambridge University Press.
- Fleischman, S. (1990). *Tense and narrativity*. Austin: Texas University Press.
- Flohr, J.W., Trevarthen, C. (2007). Music learning in childhood early developments of a musical brain and body. In Gruhn, W. and Rauscher, F. (Eds.), *Neurosciences in Music Pedagogy*, pp. 53-99.
- Fludernik, M. (1996). Towards a 'Natural' Narratology (1st ed.). Routledge.  
<https://doi.org/10.4324/9780203432501>
- Fonagy, I. (1983). *La vive voix. Essais de Psycho-Phonétique*. Paris : Bibliothèque Scientifique Payot.

- Garnica, O. (1977). *Some prosodic and paralinguistic features of speech to young children*. In C. Snow & C. Ferguson (Eds.), *Talking to young children: Language input and acquisition* (pp. 63-88). Cambridge, MA: Cambridge University Press.
- Gaskins, S. (1996). How Mayan parental theories come into play. In S. Harkness and C.M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions and consequences* (pp. 345-63). New York : The Guilford Press.
- Genette, G. (1972). *Figures: III*. Paris : Editions du Seuil.
- Gibson, J.J. (1977). The Theory of Affordances. R. Shaw, J. Bransford, eds. *Perceiving, Acting, and Knowing: towards an Ecological Psychology*. Hillsdale, NJ: Lawrence Erlbaum.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston, Houghton: Mifflin.
- Ginsburg G. P. & Kilbourne B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *J. Child Lang.* 15 221–235. 10.1017/S0305000900012344.
- Goffman, E. (1956). *The Presentation of Self in Everyday Life*. University of Edinburgh. Social Sciences Research Centre.
- Goodwin, C. (2015). *Narrative as Talk-in-Interaction*. *The Handbook of Narrative Analysis*, 195–218. doi:10.1002/9781118458204.ch10.
- Goodwin, C. (1986). Audience Diversity, Participation, and Interpretation. *Text*, 6, 283-316. <http://dx.doi.org/10.1515/text.1.1986.6.3.283>.
- Goodwin, C. (1984). *Notes on Story Structure and the Organization of Participation*. In M. Atkinson & J. Heritage (eds.), *Structures of Social Action* (pp.225-246). Cambridge: Cambridge University Press.
- Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A-Y., DeCasper, A.J. (2011). A Melodic Contour Repeatedly Experienced by Human Near-Term Fetuses Elicits a Profound Cardiac Reaction One Month after Birth. *PLoS ONE* 6(2): e17304. <https://doi.org/10.1371/journal.pone.0017304>

- Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y. and Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Developmental Science*, 14: 336-352. <https://doi.org/10.1111/j.1467-7687.2010.00978.x>
- Gratier, M. (2001). Rythmes et appartenances culturelles: Etude acoustique des échanges vocaux entre mères et bébés autochtones et migrants. Unpublished doctoral dissertation. Université René Descartes (Paris V).
- Gratier, M. (2003). Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience. *Cognitive Development*, 18 (4), 533-55.
- Gratier, M. (2007). Musicalité, style et appartenance. In M. Imberty & M. Gratier (eds), *Temps, geste et musicalité* (pp. 69-100). Paris: L'Harmattan.
- Gratier, M. (2008). Grounding in musical interaction: Evidence from jazz performances. *Musicae Scientiae Special Issue "Narrative in music and interaction"*, 71-110.
- Gratier, M. & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months, *Developmental Psychology*, 47(1):67-76.
- Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E. & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Front. Psychol.* 6:1167. doi: 10.3389/fpsyg.2015.01167
- Gratier, M. & Trevarthen, C. (2008). Musical narrative and motives for culture in mother-infant vocal interaction. *Journal of Consciousness Studies*. 15: 122-158.
- Greenfield, P. M. (1972). Playing peek-a-boo with a four-month-old: A study of the role of speech and non-speech sounds in the formation of a visual schema. *The Journal of Psychology*, 8, 287-298.
- Greenfield, P.M., Brazelton, T.B., & Childs, C. (1989). From birth to maturity in Zinacantan : Ontogenesis in cultural context. In V. Bricker and G. Gossen (Eds.), *Ethnographic encounters in Southern Mesoamerica : Celebratory essays in honour of Evon Z. Vogt* (pp. 177-216). Albany : Institute of Mesoamerican studies, State University of New York.

- Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology*, 54, 461 – 490.
- Greimas, A.J. (1969). *Éléments d'une grammaire narrative*. *L'homme*, Tome 9, pp. 71-92.
- Halbwachs, M. (1925). *Les cadres sociaux de la mémoire*. Paris: Alcan.
- Hanninen, V. (2004). A Model of narrative Circulation. In *Narrative Inquiry*, Vol 14:1, pp. 69-85.
- Haviland, L.K.M. (1978). *The social relations of work in a peasant community*. Unpublished doctoral dissertation, Harvard University, Cambridge, M.A.
- Henning, A., Striano, T., & Lieven, E. V. (2005). Maternal speech to infants at 1 and 3 months of age. *Infant behavior and Development*, 28(4), 519-536.
- Honing, H. ten Cate, C., Peretz, I. & Trehub, S.E. (2015). Without it no music: cognition, biology and evolution of musicality. *Phil. Trans. R. Soc. B* 370: 20140088.
- Huber, I. (2016). *Present-Tense Narration in Contemporary Fiction. A Narratological Overview*. London: Palgrave Macmillan.
- Hustvedt, S. (2018). Pace, space and the other in the making on fiction. *Rivista di lingue e letteratura*, Anno II no 5, pp. 23-50.
- Hymes, D. H. (ed.) (1964). *Language in culture and society: A reader in linguistics and anthropology*. New York: Harper & Row.
- Hymes, D.H. (1972) "On Communicative Competence" In: J.B. Pride and J. Holmes (eds) *Sociolinguistics. Selected Readings*. Harmondsworth: Penguin, pp. 269-293.
- Hymes, D. (1981). *In Vain I Tried to Tell You: Essays in Native American Ethnopoetics*. University of Pennsylvania Press, Philadelphia.
- Hymes, D. (1985). *Language, Memory, and Selective Performance: Cultee's "Salmon's Myth" as Twice Told to Boas*. *The Journal of American Folklore*, 98(390), 391. doi:10.2307/540365

- Hymes, D. (1996). Ethnopoetics and Sociolinguistics: Three Stories by African-American Children. In *Ethnography, Linguistics, Narrative Inequality*. London and Boston: Taylor&Francis, pp. 165-183.
- Hymes, D. (1998). When is oral narrative poetry? Generative form and its pragmatic conditions. *Pragmatics* 8:4, 475-500.
- Hymes, D. (2000). Sung epic and Native American ethnopoetics. In L. Honko (Ed.), *Textualization of Oral Epics* (pp. 291-342). Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110825848.291>
- Imberty, M. (1981). *Les écritures du temps. Semantique psychologique de la musique*. Paris: Bordas.
- Imberty, M. (2005). *La musique creuse le temps*. Paris : L'Harmattan.
- Imberty, M. (2008). Narrative, splintered temporalities and the unconscious in 20th century music. *Musicae Scientiae Special Issue: Narrative in Music and Interaction*, 129-146.
- Iyer, V. (2004). Exploding the Narrative in Jazz Improvisation. In R. O'Meally, B. Edwards & F. Griffin eds. *Uptown Conversation: The New Jazz Studies*, pp. 393-403. New York: Columbia University Press.
- Iyer, S. & Denson, H. & Lazar, N. & Oller, D. K. (2016). Volubility of the human infant: Effects of parental interaction (or lack of it). *Clinical Linguistics & Phonetics*. 30. 1-19. [10.3109/02699206.2016.1147082](https://doi.org/10.3109/02699206.2016.1147082).
- Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100, 33–72.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy. *Monographs of the Society for Research in Child Development*. 66 (2), (Serial No. 265).
- Jakobson, R. (1942). *Six Lectures on Sound and Meaning*, (Online),(<http://www.marxist.org/reference/subject/philosophy/works/ru/jakobson.htm>, retrieved on May 15, 2019).

- Jakobson, R. (1960). Linguistics and Poetics. In T. Sebeok (Ed.), *Style in Language* (pp. 350-377). Cambridge: Massachusetts Institute of Technology Press.
- Jakobson, R. (1973). *Questions de Poétique*. Paris: Editions du Seuil
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, 19(3), 627-640.
- Kärtner J., Keller H., Yovsi R. D. (2010). Mother–infant interaction during the first 3 months: the emergence of culture-specific contingency patterns. *Child Dev.* 81, 540–554.
- Kaye, K. (1980). Why we don't talk 'baby talk' to babies. *Journal of Child Language*, 7(03), 489-507.
- Keller, H. (2007). *Cultures of infancy*. Mahwah, NJ : Lawrence Erlbaum.
- Keller, H. (2003). Socialization for competence Cultural models of infancy. *Human Development*, 46, 288 – 311.
- Keller, H., Lohaus, A., Voelker, S., Cappenberg, M., & Chasiotis, A. (1999). Temporal contingency as an independent component of parenting behavior. *Child Development*, 70, 474-485.
- Keller H., Kärtner J., Borke J., Yovsi R. & Kleis A. (2005). Parenting styles and the development of the categorical self: a longitudinal study on mirror self-recognition in Cameroonian Nso and German families. *International Journal of Behavioral Development*, 29 (6), 496-504.
- Keller, H., Voelker, S., Yovsi, R.D. (2005). Conceptions of parenting in different cultural communities: The case of West African Nso and Western German women. *Social Development*, 14, 1, 158-179.

- Keller, H., Relindis, Y., Joern, B., Kartner, J., Henning, J., Papaligoura, Z. (2004). Developmental Consequences of Early Parenting Experiences: Self-Recognition and Self-Regulation in Three Cultural Communities. *Child Development*, 75 (6), 1745-1760.
- Keller, H., Borke, J., Staufenbiel, T., Yovsi, R.D., Abels, M., Papaligoura, Z., Jensen, H., Lohaus, A, Chaudhary, N., Lo, W., Su, Y. (2009). Distal and proximal parenting as alternative parenting strategies during infants' early months of life: A cross-cultural study. *International Journal of Behavioral Development*, 33(5), 412-420.
- Kessen, W., Levine, J., & Wendrich, K.A. (1979). The imitation of pitch in infants. *Infant Behavior & Development*, 2, 93-99.
- Kokkinaki, T., & Kugiumutzakis, G. (2000). Basic aspects of vocal imitation in infant-parent interaction during the first 6 months. *Journal of Reproductive and Infant Psychology*, 18(3), 173–187. <https://doi.org/10.1080/713683042>
- Koopmans-van Beinum, F.J. & van der Stelt, J.M. (1979). *Early Stages in Infant Speech Development*. Retrieved from (May 15, 2019): <http://www.fon.hum.uva.nl/archive/1979/Proc5BeinumStelt.pdf>
- Krueger, J. (2011a). Doing things with music. *Phenomenol. Cogn. Sci.* 10, 1–22. doi: 10.1007/s11097-010-9152-4
- Kruper, J.C., Uzgiris, I.C. (1987). Fathers' and mothers' speech to young infants. *Journal of Psycholinguistic Research* 16(6):597-614.
- Kugiumutzakis, G. (1985). The origin, development and function of early infant imitation. PhD. Thesis, Psychology, Uppsala University, Sweden.
- Kühl, O. (2011). *The semiotic gesture*. New perspectives on Music and Gesture. Edited by Anthony Gritten and Elaine King. Ashgate Publishing, Ltd.
- Kühl, O. (2007). *Musical semantics*. (European Semiotics: Language, Cognition and Culture. No. 7). Bern: Peter Lang.
- Labov, W. & Waletzki, J. (1967), Narrative analysis, in *Essays on the Verbal and Visual Arts*, J. helm (dir.), Seattle, University of Washington Press, pp 12-44.

- Labov, W. (1972). *Language in the Inner City*. Philadelphia: Univ. of Pennsylvania Press.
- Labov, W. (1978). *Le parler ordinaire*. Paris : Editions de Minuit.
- Labov, W. (1997). Some further steps in narrative analysis. *Journal of Narrative and Life History*, 7:395-415.
- Labov, W. (2001). Uncovering the Event Structure of Narrative. In *Georgetown Round Table*, pp. 12–44. Georgetown: Georgetown University Press.
- Labov, W. (2006). Narrative pre-construction. *Narrative Inquiry*, 16(1), 37-45.
- Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*, Chicago: University of Chicago Press.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Langer S. K. (1942). *Philosophy in a new key; a study in the symbolism of reason rite and art*. Harvard University Press.
- Langer, S. K. (1953). *Feeling and Form*. New York: Scribner's.
- Langer, S. K. (1966). The Cultural Importance of the Arts. *Journal of Aesthetic Education*, 1(1), pp. 5 – 12.
- Langer, S. K. (1988). *Mind: An Essay on Human Feeling (Abridged Edition)*. Baltimore, MD: Johns Hopkins University Press.
- Lavelli, M., & Fogel, A. (2005). Developmental changes in the relationship between the infant's attention and emotion during early face-to-face communication: the 2-month transition. *Developmental Psychology*, 41(1), 265.
- Lavelli, M. & Fogel, A. (2002). Developmental changes in mother-infant face-to-face communication: birth to 3 months. *Developmental Psychology*, 38 (2), 288-305.
- Lee, D. & Schögler, B. (2009). Tau in Musical Expression. In S. Malloch & C. Trevarthen (Eds.) *Communicative Musicality: Exploring the Basis of Human Companionship* (pp.83-104). Oxford: Oxford University Press.
- Lerdahl, F. & Jackendoff, R.S. (1983). *A Generative Theory of Tonal Music*. MIT Press.



- LeVine, R.A. (1970). Cross-cultural study in child psychology. In P.H. Mussen (ed.), *Carmichael's Manual of Child Psychology*. New York: John Wiley.
- LeVine, R.A., Dixon, S., Levine, S., Richman, A., Liederman, P.H., Keefer, C.H., & Brazelton, T.B. (1996). *Child care and culture: Lessons from Africa*. Cambridge: Cambridge University Press.
- Levinson, J. (2004). Music as Narrative and Music as Drama. *Mind and Language*, 19(4), 428–441.
- Levinson, S.J. (2016). Turn-taking in Human Communication. Origins and Implications for Language Processing. *Trends Cogn Sci.*, 20(1):6-14.
- Lévi-Strauss, C. (1958). *Anthropologie structurale*. Plon, Paris.
- Lewis, G.E. (1996). "Improvised Music Since 1950: Afrological and Eurological Forms." *Black Music Research Journal* 16-1, pp. 91-119.
- Lynch, M. P., Oller, D. K., Steffens, M. L., & Buder, E. H. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychobiology*, 28(1), 3–25.
- Malloch, S.N. (1999). Mothers and infants and communicative musicality. *Musicae Scientiae, Special Issue*, 1999-2000, 29-57.
- Malloch, S.N. & Trevarthen, C. (Eds) (2009). *Communicative Musicality: Exploring the Basis of Human Companionship*. Oxford: Oxford University Press.
- Mampe, B., Friederici, A.D., Christophe, A., Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology* 19, 1994-1997.
- Margulis, E.H., Wong, P.C.M., Turnbull, C., Kubit, B.M., & McAuley, J.D. (2022). Narratives Imagined in Response to Instrumental Music Reveal Culture-Bounded Intersubjectivity. *Proceedings of the National Academy of Sciences*, 119(4), e2110406119. <https://doi.org/10.1073/pnas.2110406119>
- Markus, H. & Nurius, P. (1986). Possible Selves. *American Psychologist*, 41, 954-969.

- Masataka, N. (2017). Neurodiversity, Giftedness, and Aesthetic Perceptual Judgment of Music in Children with Autism. *Frontiers in Psychology*, 8, 1595. 10.3389/fpsyg.2017.01595.
- Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental science*, 9(1), 46-50.
- Masataka, N. & Bloom, K. (1994). Acoustic properties that determine adults' preferences for 3-month-old infant vocalizations. *Infant Behavior & Development*, 17(4), 461-464.
- Maus, F.E. (1991). Music as narrative. In *Indiana Theory Review* Vol.12 Spring and Fall, pp. 1-34.
- Mazokopaki, K., & Kugiumutzakis, G. (2009). Infant rhythms: Expressions of musical companionship. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 185-208). Oxford: Oxford University Press.
- McGowan, T., & Delafield-Butt, J. (2022). Narrative as co-regulation: A review of embodied narrative in infant development. *Infant behavior & development*, 68, 101747. Advance online publication. <https://doi.org/10.1016/j.infbeh.2022.101747>
- Mead, G.H. (1934). The Vocal Gesture and the Significant Symbol. In *Mind Self and Society from the Standpoint of a Social Behaviorist*. Edited by Charles W. Morris. Chicago: University of Chicago, 61-68.
- Mehler, J., Bertoncini, J., Barrière, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception*, 7(5), 491-497. <https://doi.org/10.1068/p070491>
- Meltzoff, A.N. & Moore, M.H. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75-78.
- Meltzoff, A.N. & Moore, M.K. (1992). Early imitation within a functional framework: the importance of personal identity, movement and development. *Infant Behavior and Development*, 15, 479-505.

- Meltzoff, A.N. & Moore, M.K. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, 17, 83-99.
- Merker, B. (2009). Ritual foundations of human uniqueness. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 45–59). Oxford University Press.
- Merleau-Ponty, M. (1962). *Phenomenology of perception*. London: Routledge.
- Messinger, D., & Fogel, A. (2007). The interactive development of social smiling. *Advances in Child Development and Behaviour*, 35, 328-366.
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: Chicago University Press.
- Miall, D. S., & Dissanayake, E. (2003). The poetics of babytalk. *Human Nature*, 14, 337-364
- Miell, D., MacDonald, R. and Hargreaves, D.J. (Editors) (2005). *Musical Communication*. New York: Oxford University Press.
- Monson, I. (1996). *Saying something: Jazz improvisation and interaction*. Chicago: The University of Chicago Press.
- Murray, L., & Trevarthen, C (1985). Emotional regulation of interactions between two-month-olds and their mothers. In Field, T M & Fox, N A (Eds.) *Social Perception in Infants*. (pp. 177-197). Norwood, N J: Ablex.
- Nakata, T. & Trehub, S. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain*. 21. 10.1037/h0094003.
- Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago, IL, US: University of Chicago Press.
- Nattiez, J-J. (2013). The Narrativization of Music. Music: Narrative or Proto-Narrative? *Human & Social Studies. Research and Practice*. 2. 10.2478/hssr-2013-0004.
- Néda, Z., Ravasz, E., Vicsek, T., Bréchet, Y., & Barabasi, A.L. (2000). Physics of the rhythmic applause. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 61 6 Pt B, 6987-92.

- Nelson, K. (2006). *Narratives from the Crib*. Cambridge, MA: Harvard University Press.
- Nelson, K. (1993). The Psychological and Social Origins of Autobiographical Memory. *Psychological Science*, Volume: 4 Issue: 1, page(s): 7-14.
- Nettle, B. (1983). *The Study of Ethnomusicology: Thirty-one Issues and Concepts*. University of Illinois Press. Urbana and Chicago.
- Norman-Haignere, S.V., Kanwisher, N.G., and McDermott, J.H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron* 88, 1281–1296.
- Norman-Haignere, S. V., Feather, J., Boebinger, D., Brunner, P., Ritaccio, A., McDermott, J. H., Schalk, G., & Kanwisher, N. (2022a). A neural population selective for song in human auditory cortex. *Current biology : CB*, S0960-9822(22)00131-2. Advance online publication. <https://doi.org/10.1016/j.cub.2022.01.069>
- Norman-Haignere, S.V., Long, L.K., Devinsky, O., Doyle, W., Irobunda, I., Merricks, E.M., Feldstein, N.A., McKhann, G.M., Schevon, C.A., Flinker, A., (2022b). Multiscale temporal integration organizes hierarchical computation in human auditory cortex. *Nat Hum Behav* 6, 455–469. <https://doi.org/10.1038/s41562-021-01261-y>
- Ochs, E. & Capps, L. (1996). Narrating the Self. *Annual Review of Anthropology*, Vol. 25, pp. 19-43.
- Ochs, E. (1997). Narrative. In Teun A. Van Dijk Eds, *Discourse as structure and process*, pp 185-208.
- Ochs, E. & Capps, L. (2002). *Living Narrative. Creating Lives in Everyday Storytelling*. Harvard: Harvard University Press.
- Ochs, E. & Shohet, M. (2006). “The Cultural Structuring of Mealtime Socialization.” In *New Directions in Child and Adolescent Development Series Number 11: Family Mealtime as a Context of Development and Socialization*. San Francisco: Jossey-Bass, pp. 35-50.
- Ochs, E. (2011). “Narrative in Everyday Life” (Revised).” In *Discourse Studies: A Multidisciplinary Introduction*. Teun A. van Dijk (ed.) London: Sage.

- Oller, D.K. (1986). Metaphonology and infant vocalizations. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 21-36). Basingstroke, Hampshire, United Kingdom: Macmillan.
- Oller, D.K. & Lynch, M.P. (1992). Infant vocalizations and innovations in infraphonology: Toward a broader theory of vocal development and disorders. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development* (pp. 509-536). Parkton, MD: York Press.
- Oller, D. K. (1980). The emergence of the sounds of speech in infancy, in *Child Phonology*, Vol. 1. Production, eds Yeni-Komshian G. H., Kavanaugh J. F., Ferguson C. A., editors. (New York, NY: Academic Press; ), 93–112.
- Oller, D.K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum and Associates.
- Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16(4), 728–741. <https://doi.org/10.1037/0096-1523.16.4.728>
- Papoušek, M., & Papoušek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. *Advances in Infancy Research*, 1, 163-224.
- Papousek, H., Papousek, M., & Bornstein, M. H. (1985). The naturalistic environment of young infants: On the significance of homogeneity and variability in parental speech. In T. Field, & N. Fox (Eds.), *Social perception in infants* (pp. 269–297). Norwood, NJ: Ablex.
- Papoušek, M. & Papoušek, H. (1989a). Forms and functions of vocal matching in interactions between mothers and their pre-canonical infants. *First Language*, 9, 137-158.
- Papousek, H., and Papousek, M. (1989b). Intuitive parenting: Aspects related to educational psychology. In B. Hopkins, M.-G. Pecheux, and H. Papousek (Eds.), *Infancy and education: Psychological considerations*. *European Journal of Psychology of Education*, 4 (2, Special Issue), 201–210.

- Papoušek, H., & Papoušek, M. (2002). Intuitive parenting. *Handbook of Parenting Volume 2 Biology and Ecology of Parenting*, 182-203.
- Papousek, M., Papousek, H., & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. *Journal of Psycholinguistic Research*, 16(5), 491-516.
- Papousek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. *Musical beginnings: Origins and Developments of Musical Competence*, 88-112.
- Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology*, 2, 142. <https://www.frontiersin.org/article/10.3389/fpsyg.2011.00142>
- Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.
- Pêcheux, M. G., Labrell, F., & Pistorio, M. (1993). What do parents talk about to infants?. *Early Development and Parenting*, 2(2), 89-97.
- Peretz, I., Vuvan, D., Lagrois, M. É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370 (1664), 20140090. <https://doi.org/10.1098/rstb.2014.0090>
- Porges, S.W. (1992). Vagal tone: a physiologic marker of stress vulnerability. *Pediatrics*, 90: 498–504.
- Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation*. The Norton series on interpersonal neurobiology. New York, NY, US: WW Norton & Co.
- Propp, V. (1928/2003). *Morphology of the Folktale*. Austin: University of Texas Press.
- Rabain-Jamin J., Sabeau-Jouannet, E. (1997). Maternal speech to 4-month-old infants in two cultures: Wolof and French. *Int. J. Behav. Dev.* 20, 425–451.

- Rączaszek-Leonardi, J., Rossmanith, N., Nomikou, I. & Rohlfing, K. (2019). Levels of Coordination in Early Semantic Development. *Psychology of Language and Communication*, 23(1) 212-237. <https://doi.org/10.2478/plc-2019-0010>
- Reese, E., Haden, C.A., Baker-Ward, L., Bauer, P., Fivush, R. & Ornstein, P.A. (2011). Coherence of Personal Narratives across the Lifespan: A Multidimensional Model and Coding Method. *J Cogn Dev.*, 12(4):424-462.
- Reybrouck, M., Vuust, P., & Brattico, E. (2018). Brain Connectivity Networks and the Aesthetic Experience of Music. *Brain sciences*, 8(6), 107. <https://doi.org/10.3390/brainsci8060107>
- Reybrouck, M. (2005). Body, mind and music: musical semantics between experiential cognition and cognitive economy. *Trans: Transcultural Music Review*, 9. Retrieved May 4, 2019, from: [https://limo.libis.be/primoexplore/fulldisplay?docid=LIRIAS1797519&context=L&vid=Lirias&search\\_scope=Lirias&tab=default\\_tab&lang=en\\_US&fromSitemap=1](https://limo.libis.be/primoexplore/fulldisplay?docid=LIRIAS1797519&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US&fromSitemap=1)
- Ricoeur, P. (1983). *Temps et Récit*. Paris : Seuil.
- Ricoeur, P. (1990). Le soi et l'identité narrative. In *Soi-même comme un autre*, pp 190-192.
- Rizzolatti, G. & Gallese, V. (2003). Mirror neurons. In: Nadel L, (ed.) *Encyclopedia of Cognitive Science*. London: Nature Publishing Group, pp.37-42.
- Rohlfing, K. J., Leonardi, G., Nomikou, I., Rączaszek-Leonardi, J. & Hüllermeier, E. (2020). Multimodal Turn-Taking: Motivations, Methodological Challenges, and Novel Approaches. *IEEE Transactions on Cognitive and Developmental Systems*, Volume 12, No 2, pp. 260-271. doi: 10.1109/TCDS.2019.2892991.
- Rossmanith, N., Costall, A., Reichelt, A. F., López, B., & Reddy, V. (2014). Jointly structuring triadic spaces of meaning and action: book sharing from 3 months on. *Frontiers in Psychology*, 5, 1390. <https://doi.org/10.3389/fpsyg.2014.01390>
- Rudrum, D. (2006). On the very idea of a definition of narrative: A reply to Marie-Laure Ryan. *Narrative*, Volume 14, No 2, The Ohio State University, pp. 188-196.

- Sacks H., Schegloff, E. A., Jefferson G. (1974). A simplest systematics for the organization of turn taking in conversation. *Language*, 50, 696–735.
- Samson, J. (1974). Schoenberg's 'Atonal' Music. In *Tempo*, No. 109 (Jun., 1974), pp. 16-25.
- Sapir, E. (1929). A study in phonetic symbolism. *J. Exp. Psychol.* 12, 225–239.  
doi:10.1037/h0070931
- Saussure, F. de (1972/1916). *Cours de linguistique générale*, édition critique préparée par Tullio De Mauro. Paris : Payot.
- Savage, P.E. , Brown, S. , Sakai, E. , Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112 (29) 8987-8992; DOI:10.1073/pnas.1414495112.
- Sawyer, R. K. (1992). Improvisational creativity: An analysis of jazz performance. *Creativity Research Journal* 5, 253-263.
- Sawyer, R.K. (2002). Improvisation and Narrative. *Narrative Inquiry*, Volume 12, Number 2, 2002, pp. 319-349(31).
- Schank, R.C. (1975). *Conceptual Information Processing*. New York: Elsevier.
- Schank, R.C. & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*. Hillsdale , NJ: Earlbaum Assoc.
- Schegloff, E. A. (1997). 'Narrative Analysis' Thirty Years Later. *Journal of Narrative and Life History*, 7 (1–4): 97–106.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143-165.
- Scherer, K. R. (2003). Vocal communication of emotion. *Speech and Communication*, 40(1–2), 227–256.
- Schieffelin, B., & Ochs, E. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. Shweder, & R. Levine (Eds.), *Culture Theory: Essays on Mind, Self and Emotion* (pp. 276-320). Cambridge University Press.



- Schieffelin, B.B. (1990). *The give and take of everyday life: Language Socialization of Kaluli children*. Cambridge: Cambridge University Press.
- Schieffelin, B.B., Ochs, E. (1998). A Cultural Perspective on the Transition from Prelinguistic to Linguistic Communication. In *Cultural Worlds of Early Childhood*, 115-131. Edited by Woodhead, Faulkner and Littleton. Routledge.
- Schoenberg, A. (1967). *Fundamentals of Musical Composition*. London: Faber and Faber.
- Scholes, R. (1980). Language, Narrative, and Anti-Narrative. *Critical Inquiry*, 7, no. 1: 204-212.
- Schögler, B. (2002). *The pulse of communication in improvised jazz duets*. Unpublished doctoral dissertation, The University of Edinburgh, U.K.
- Schögler, B. (1998). Music as a Tool in Communications Research, *Nordisk Tidsskrift for Musikterapi*, 7:1, 40-49, DOI: 10.1080/08098139809477919
- Sidnell, J. & Stivers, T. (eds.) (2012). *The Handbook of Conversation Analysis*. Wiley-Blackwell.
- Smith, N.A. & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13(4), 410-420.
- Snow, C.E. (1972). Mothers' Speech to Children Learning Language. *Child Development* Vol. 43, No. 2, pp. 549-565.
- Snow, C.E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, 4(1), 1-22.
- Snow, C.E., van Eeden, R. & Muysken, P.C. (1981). The Interactional Origins of Foreigner Talk: Municipal Employees and Foreign Workers. *International Journal of the Sociology of Language*, 28, pp. 81-91.
- Soderstrom, M., Bolossom, M., Foygel, R., Morgan, J.L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language* 35(4):869-902.

- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501–532.
- Stark R. E. (1980). Stages of speech development in the first year of life, in *Child Phonology* Vol. 1. Production, eds Yeni-Komshian G. H., Kavanaugh J. F., Ferguson C. A., editors. (New York, NY: Academic Press; ), 73–92.
- Stark, R.E., Bernstein, L.E., Demorest, M.E. (1993). Vocal Communication in the First 18 Months of Life. *Journal of Speech and Hearing Research*, Volume 36, 548-558.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10587–10592. doi: 10.1073/pnas.0903 616106
- Stern, D. N. (1982). Some interactive functions of rhythm changes between mother and infant. In M. Davis (Ed), *Interaction rhythms: Periodicity in communicative behavior* (pp.101-117). New York: Human Sciences Press.
- Stern, D. N. (1992). L’enveloppe prénarrative: Vers une unité fondamentale d’expérience permettant d’explorer la réalité psychique du bébé. *Revue internationale de psychopathologie*, 6, 13-63.
- Stern, D. N. (1993). The role of feelings for an interpersonal self. In U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge* (pp. 205-215). New York: Cambridge University Press.
- Stern, D. N. (1995). *The Motherhood Constellation. A Unified View of Parent-Infant Psychotherapy*. New York: Basic Books.
- Stern, D. N. (1998). The process of therapeutic change involving implicit knowledge: Some implications of developmental observations for adult psychotherapy. *Infant Mental Health Journal*, 19(3), 300-308.
- Stern, D. N. (1999). Vitality contours: The temporal contour of feelings as a basic unit for constructing the infant’s social experience. In Rochat, P. (ed), *Early social cognition: Understanding others in the first months of life* (pp. 67-90). Mahwah, N.J.: Erlbaum.

- Stern, D. N. (2000). *The interpersonal world of the infant: A view from psychoanalysis and development psychology*. (Originally published in 1985). Basic Books: New York.
- Stern, D. N. (2004). *The present moment in psychotherapy and everyday life*. New York: Norton & Company.
- Stern, D. N. (2010). *Forms of Vitality. Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy, and Development*. Oxford: Oxford University Press.
- Stern, D. N., Beebe, B., Jaffe, J., and Bennett, S. L. (1977). The infant's stimulus world during social interaction: a study of caregiver behaviors with particular reference to repetition timing. In H. R. Schaffer (Ed.), *Studies in Mother–Infant Interaction* (pp.177– 202). New York, NY: Academic Press.
- Stern, D.N, & Gibbon, J. (1979). Temporal expectancies of social behaviours in mother-infant play. In E.B. Thoman (Ed.), *Origins of the Infant's Social Responsiveness* (pp. 409–429). Hillsdale, NJ : Lawrence Erlbaum Associates.
- Stern, D. N., Jaffe, J., Beebe, B., and Bennett, S. L. (1975). Vocalizing in unison and in alternation: two modes of communication within the mother-infant dyad. *Ann. N. Y. Acad. Sci.* 263, 89–100.
- Stern, D , Spieker, S, Barnett, R , & MacKain, K. (1983). The prosody of maternal speech Infant age and context related changes *Journal of Child Language* 10, 1-15.
- Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, 18(5), 727-735.
- Stork, H. E. (1986). *Enfances indiennes : Etude de psychologie transculturelle et comparée du jeune enfant*. Paris: Le Centurion.
- Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists*. New York: Cambridge University Press.
- Tarasti, E. (1988). "On the modalities and narrativity in music", in: *Essays on the philosophy of music* (Acta Philosophica Fennica 43.), V.Rantala – L. Rowell –E. Tarasti (eds.) 110-131. Helsinki: Akateeminen Kirjakauppa.

- Tarasti, E. (1994). *A Theory of Musical Semiotics*. Bloomington: Indiana University Press.
- Toda, S., Fogel, A., & Kawai, M. (1990). Maternal speech to three-month-old infants in the United States and Japan. *Journal of Child Language*, 17(2), 279-294.
- Todorov, T. (1966). Les catégories du récit littéraire. In: *Communications*, 8, Recherches sémiologiques : l'analyse structurale du récit. pp. 125-151.
- Todorov, T. (1968). La grammaire du récit. In: *Langages*, no 12, pp. 94-102.
- Todorov, T. (1971). The Two Principles of Narrative. *Diacritics*, Vol. 1, No. 1, pp. 37-44.
- Todorov, T. (1978). *Symbolisme et interprétation*. Seuil, Paris.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Trehub, S. E., Ghazban, N., & Corbeil, M. (2015). Musical affect regulation in infancy. *Annals of the New York Academy of Sciences*, 1337, 186-192.
- Trehub, S. E., Plantinga, J., Brcic, J., & Nowicki, M. (2013). Cross-modal signatures in maternal speech and singing. *Frontiers in Psychology*, 4, Article 811. <https://doi.org/10.3389/fpsyg.2013.00811>
- Trehub, S. & Trainor, L. (1998). Singing to infants: Lullabies and play songs. *Advances in Infancy Research*. 12. 43-77.
- Trehub, S.E., Trainor, L.J., & Unyk, A.M. (1993). Music and speech processing in the first year of life. *Advances in Child Development and Behaviour*, 24, 1-35.
- Trehub, S.E., Unyk, A.M. & Trainer, L.J. (1993). Maternal Singing in Cross-Cultural Perspective. *Infant Behavior and Development*, 16, 285-295.
- Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, 21(1), 77-88.
- Trainor, L.J., Clark, E.D., Huntley, A., & Adams, B.A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior & Development*, 20 (3), 383-396.

- Turner, M. (1996). *The Literary Mind: The origins of thought and language*. Oxford: Oxford University Press.
- Trevarthen, C. (1974). Conversations with a 2-month-old. *New scientist*, 230-235.
- Trevarthen, C. (1977). Descriptive analyses of infant communicative behavior. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 227-270). London: Academic Press.
- Trevarthen, C., and Hubley, P. (1978). "Secondary intersubjectivity: confidence, confiding and acts of meaning in the first year," in *Action, Gesture and Symbol: The Emergence of the Language*, ed. A. Lock (London: Academic Press), 183–229.
- Trevarthen, C. (1979). "Communication and cooperation in early infancy: a description of primary intersubjectivity," in *Before Speech: The Beginning of Interpersonal Communication*, ed. M. Bullowa (Cambridge: Cambridge University Press), 321–347.
- Trevarthen, C. (1989). Motives for culture in young children - their natural development through communication. In, W. Koch (Ed.) *The nature of culture*. (pp. 80-119) Bochum: Brockmeyer.
- Trevarthen, C. (1993). The self born in intersubjectivity: An infant communicating. In U. Neisser (Ed.), *The perceived self: Ecological and interpersonal sources of self-knowledge*. New York: Cambridge University Press, 121-173.
- Trevarthen, C. (1996). Lateral asymmetries in infancy: Implications for the development of the hemispheres. *Neuroscience and Biobehavioral Reviews*, 20. 571-586.
- Trevarthen, C. (1998). The concept of foundations of infant intersubjectivity. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 15-47). Cambridge, UK: Cambridge University Press.
- Trevarthen, C. (2000). Musicality and intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae Special Issue 1999-2000*, 155-211.

- Trevarthen, C. (2001). *The neurobiology of early communication: intersubjective regulations in human brain development*. In: Handbook on Brain and Behavior in Human Development, Kalverboer AF, Gramsbergen A, eds. Dordrecht, The Netherlands: Kluwer, pp. 841-882
- Trevarthen, C. (2008). The musical art of infant conversation: Narrating in the time of sympathetic experience, without rational interpretation, before words. *Musicae Scientiae*, Volume: 12 Issue: 1\_suppl, page(s): 15-46.
- Trevarthen, C. & Malloch, S. (2002). Musicality and music before three: Human vitality and invention shared with pride. *Zero to Three*, 23(1), 10-18.
- Tronick, E., Als, H., Adamson, L., Wise, S., and Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *J. Am. Acad. Child Psychiatry* 17, 1–13. doi: 10.1016/S00027138(09)62273-1
- Ullal-Gupta, S., Vanden Bosch der Nederlanden, C.M., Tichko, P., Lahav, A., & Hannon, E.E. (2013). Linking prenatal experience to the emerging musical mind. *Front. Syst. Neurosci.*7:48. doi:10.3389/fnsys.2013.00048.
- Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and Simplicity: A Cross-Cultural Perspective. *Psychology of Music*, 20(1), 15–28. <https://doi.org/10.1177/0305735692201002>
- Van Dijk, T.E. (1975). Action, Action Description and Narrative. *New Literary History*, 6, p. 273-294.
- Van Dijk, T. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.
- Van Egeren, L. A., Barratt, M. S., and Roach, M. A. (2001). Mother–infant responsiveness: timing, mutual regulation, and interactional context. *Dev. Psychol.*37,684–697.doi:10.1037/0012-1649.37.5.684
- Van Puyvelde, M., Loots, G., Vanfleteren, P., Meys, J., Simcock, D., & Pattyn, N. (2014). Do you hear the same? Cardiorespiratory responses between mothers and infants during

tonal and atonal music. *PloS one*, 9(9), e106920.  
<https://doi.org/10.1371/journal.pone.0106920>

Weinberg, M.K. & Tronick, E.Z (1994). Beyond the face : An empirical study of infant affective configurations of facial, vocal, gestural, and regulatory behaviors. *Child Development*, 65, 1503-1515.

Williams, J. A., Margulis, E. H., Nastase, S. A., Chen, J., Hasson, U., Norman, K. A., & Baldassano, C. (2022). High-Order Areas and Auditory Cortex Both Represent the High-Level Event Structure of Music. *Journal of cognitive neuroscience*, 34(4), 699–714.  
[https://doi.org/10.1162/jocn\\_a\\_01815](https://doi.org/10.1162/jocn_a_01815)

Yoo, H., Bowman, D.A. and Oller, D.K. (2018). The Origin of Protoconversation: An Examination of Caregiver Responses to Cry and Speech-Like Vocalizations. *Front. Psychol.* 9:1510. doi: 10.3389/fpsyg.2018.01510

Zentner, M. & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proc Natl Acad Sci U S A*.107(13): 5768-73.

