Universidade do Minho
Escola de Psicologia

Vladimiro Gomes Lourenço

Microdynamics of Interpersonal Coordination: A Microanalytic Perspective on the Development of Turn-Taking

abril de 2022

UMinho|2022
Vladimiro Gomes Lourenço
Microdynamics of Interpersonal Coordination:
A Microanalytic Perspective on the Development of Turn-Taking

FCT Fundação para a Ciência e a Tecnologia

Cofinanciado por:

COMPETE 2020

PORTUGAL 2020

UNIÃO EUROPEIA
Fundo Social Europeu

Universidade do Minho
Escola de Psicologia

Vladimiro Gomes Lourenço

Microdynamics of Interpersonal Coordination: A Microanalytic Perspective on the Development of Turn-Taking

Tese de Doutoramento
Doutoramento em Psicologia Básica

Trabalho efetuado sob a orientação do
Professor Doutor Alfredo Manuel Feliciano Pereira
e da
Professora Doutora Joana Fernandes Pereira Coutinho

abril de 2022

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Acknowledgements

I would like to thank Portuguese Foundation for Science and Technology (FCT) for funding this research with a doctoral grant (reference PD/BD/128460/2017).

I thank my supervisors Alfredo F. Pereira and Joana Coutinho for all the contributions in this dissertation, and for actively and supportively sharing this journey with me.

I am also grateful to Jorge A. Santos for the earlier contributions to this project and for letting me be a part of the Laboratory of Visualization and Perception.

I thank Ana Paula Soares for her support and advices, both as an element of the PhD Committee and as Director of the Doctoral Program in Basic Psychology.

I would also like to thank to Adriana Sampaio for providing the database for one of the studies in this dissertation.

I would like to thank to all the participants in the studies in this dissertation.

I would like to thank Juliana Serra for the multiple roles played in this dissertation – helping in the experiment design, in participants recruitment, in data collection, as a participant experimenter, and in data segmentation. Above all, for being a major source of support throughout this journey.

I am also grateful to Isabel Lisboa for helping me from the very first steps into this journey to the final ones.

I thank Daniel Basso, Sandra Queirós, Filipe Araújo, Ana Alexandra Moura, and Joana Ribeiro, for each contributions to the studies in this dissertation.

I thank my friends for all the shared moments along the way.

Finally, I would like to thank my family. My mother for all the worrying and caring. My father for have letting me take care of him as best I could. My brother for always being there when I need. And my sisters for all the care and support they provide to the family.

Thank you all.

Statement Of Integrity

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Microdinâmicas da Coordenação Interpessoal: Uma Perspectiva Microanalítica sobre o Desenvolvimento da Tomada-de-Vez

Resumo

A tomada-de-vez corresponde ao padrão de alternância que organiza a conversação humana e determina que se fale um de cada vez, evitando sobreposições e silêncios prolongado entre turnos. As transições de turno ocorrem a velocidades impressionantes que desafiam até as capacidades mais básicas de produção linguística. Apesar disso, existem evidências de que até os bebés conseguem manter padrões de alternância semelhantes nas suas trocas vocais, muito antes do desenvolvimento da linguagem. Nesta dissertação, propomo-nos a seguir a trajetória desenvolvimental da duração das transições de turno, em bebés (7, 12 meses) e crianças (3-5 anos). Iremos usar uma abordagem microanalítica ao examinar um espectro alargado de vocalizações e definindo as transições de turno em termos de estados diádicos, derivados das interações vocais dos participantes. No Capítulo I, iremos primeiro explorar os aspetos temporais da coordenação interpessoal e os desafios metodológicos ao estudo das interações sociais. Prosseguiremos com a descrição das propriedades e modelos da tomada-de-vez, assim como as evidências evolucionárias, interculturais e desenvolvimentais dos seus alicerces. Iremos, depois, rever a mais recente investigação acerca do desenvolvimento da duração das transições de turno. No Capítulo II, iremos apresentar o Estudo 1, onde investigamos a trajetória desenvolvimental, aos 7 e aos 12 meses, das durações das sobreposições e silêncios entre turnos, em interações mãe bebé face-a-face e orientadas aos objetos. No Capítulo III, vamos reportar o Estudo 2, onde examinamos a trajetória desenvolvimental da tomada-de-vez, dos 3 aos 5 anos, em interações mãe criança de brincadeira livre. No Capítulo IV, apresentamos o Estudo 3, uma visão metodológica acerca da microanálise. Finalmente, no Capítulo V, revemos os principais resultados dos estudos nesta dissertação, refletimos sobre algumas das suas limitações e elaboramos planos para investigação futura.

Palavras-chave: duração da transição de turno, microanálise, coordenação interpessoal, tomada-de-vez, trajetórias desenvolvimentais.

Microdynamics of Interpersonal Coordination: A Microanalytic Perspective on the Development of Turn-Taking

Abstract

Turn-taking is the recurrent pattern of alternation that organizes human conversation, determining that each partner speaks mostly one at a time, while avoiding superpositions and prolonged silences between turns. These turn-transitions occur at impressive timings that challenge our most basic linguistic production abilities. In spite of that, there is evidence that even infants engage in similar patterns of vocal exchanges, long before language acquisition. In this dissertation we propose to follow the developmental trajectory of turn-transition durations in infancy (7, 12 months) and in childhood (3-5 years). We will use a microanalytic approach by focusing on the entire spectrum of vocalizations and defining turn-transitions from dyadic states of interactional vocal behavior. In Chapter I, we will first explore the temporal aspects of interpersonal coordination and the methodological challenges of investigating interactional behavior. We will proceed by describing the properties, models and the evolutionary, cross-cultural and developmental foundations of turn-taking. We will then review the latest research on the development of turn-transition duration. In Chapter II, we will present Study 1, where we investigate the developmental trajectory of gap and overlap durations, from 7 to 12 months, in mother-infant object-oriented and face-to-face interactions. In Chapter III, we will report Study 2, where we examine the developmental trajectories of turn-taking, with 3- to 5-years-old children, in mother-children free-play interactions. In Chapter IV, we will present Study 3, a methodological overview of microanalysis. Finally, in Chapter V, we will review the main findings of the empirical studies in this dissertation, reflect on some of its limitations, and elaborate on future directions we will like to explore.

Keywords: developmental trajectories, interpersonal coordination, microanalysis, turn-taking, turn-transition duration.

Table of Contents

Index of Abbreviations

ARIMA - Auto Regressive Integrated Moving Average

AVTA - Automatic Vocal Transaction Analyzer

CCF - Cross-Correlation Function

DV - Dependent Variable

EEG - Electroencephalography

fMRI - Functional Magnetic Resonance Imaging

fNIRS - Functional Near-Infrared Stereoscopy

FTO - Floor-Transfer Offset

IMUs - Inertial Measurement Units

ISS - Interruptive Simultaneous Speech

MEA - Motion Energy Analysis

MEG - Magnetoencephalography

NHI - Natural History of an Interview

NSS - Non-Interruptive Simultaneous Speech

P - Pause

SP - Switching Pause

V – Vocalization

List of Figures

List of Tables

CHAPTER I – GENERAL INTRODUCTION

General Introduction

Turn-taking is the predominant organization of human conversation (Levinson, 2006). It is characterized by a recurrent pattern of alternation between conversational partners, where each partner speaks mostly one at a time, avoiding superpositions and prolonged silences between turns (Sacks et al., 1974). This property of conversational turn-taking, known as the minimal-gap minimal-overlap phenomenon, acts as a glue that holds turns together (Casillas, 2014), determining that turn-transitions occur at impressive speeds (~250ms), requiring a great deal of coordination between partners which nonetheless appears effortless and natural (Levinson & Torreira, 2015).

This type of coordinative effort is deeply ingrained in the social matrix of our species. Interpersonal coordination is often regarded as the "social glue" that brings people together and shape our own capacity to be with one another. It is most essential as a platform for joint action and collaboration between humans (Levinson, 2006); and there is robust evidence that interpersonal coordination promotes positive intrapersonal and interpersonal outcomes – by reducing anxiety, increasing feelings of harmony between interactional partners, and promoting prosocial behavior (Vicaria & Dickens, 2016). Developmental research has also shown that mother-infant coordination, as early as 3 to 4-months-old, can even predict attachment style and cognitive development at later ages (Beebe et al., 2010; Beebe & Steele, 2016; Feldman et al., 1996; Feldman & Greenbaum, 1997; Jaffe et al., 2001).

Interpersonal Coordination

Interpersonal coordination can be defined as "*the degree to which the behaviors in an interaction are non-random, patterned or synchronized in both timing and form*" (Bernieri & Rosenthal, 1991; pp. 402-403). Most commonly, as highlighted by Paxton's (2015) scientometric analysis, interpersonal coordination has become an umbrella for several related terms (e.g., accommodation, adaptation, alignment, coordination, contagion, contingency, convergence, coupling, entrainment, linkage, matching, mimicry, mirroring, synergy, synchrony) sometimes used interchangeably, leading to much conceptual ambiguity (Burgoon et al., 1995; Paxton, 2015). This is nonetheless a result of the diversity in the study of interpersonal coordination, where different research traditions operationalize interpersonal coordination at different levels of abstraction and analysis (Paxton, 2015).

Conceptual Aspects of Interpersonal Coordination

To the purpose of this dissertation, we are particularly interested in the temporal aspects of interpersonal coordination. Among them, Lewkowicz's (1989) points to a distinction between temporal relations (synchrony, contingency); temporal features (duration, rhythm); and temporal predictability (anticipation on basis of temporal regularity). In relation to the study of turn-taking, all of these dimensions are relevant, and particularly duration and predictability will be further explored in other sections. In terms of our conceptualization of interpersonal coordination, we are most definitely interested in the temporal relations between the behavior of a partner and the behavior of another – sometimes described as interactive contingency (Burgoon et al., 1995; Jaffe et al., 2001), more often equated to one of the oldest terms used in interpersonal coordination research: synchrony (Paxton, 2015).

Interpersonal Synchrony

The concept of interpersonal synchrony began to be explored in the 1960's from the microanalytical works of Condon (1970; Condon & Ogston, 1966, 1967) and Kendon (1970) and was demonstrated in the almost simultaneous changes in infant movement (type, joints) to the rhythm of their mother's vocalizations (Condon & Sander, 1974). The validity of interactional synchrony as a genuine phenomenon was then challenged by McDowall (1978a, 1978b), who argued against the replicability of the methodology and considered the phenomenon as a statistical artifact. This challenges were later found inconclusive and irrelevant to previous research, given noticeable differences in the application of the methodology (Gatewood & Rosenwein, 1981). Interpersonal synchrony as since then been recognized as a relevant and multidisciplinary phenomenon (Delaherche et al., 2012).

In their massive effort to review and re-conceptualize the dynamics of interpersonal adaptation, Burgoon et al. (1995) highlighted a relevant distinction within interactional synchrony: between simultaneous and concactenous interpersonal behavior. While the former corresponds to the concept of interactional partners performing actions at the same time, the later denotes another form of interpersonal temporal organization, where interaction partners coordinate their behavior sequentially, by taking turns (Burgoon et al., 1995). Conversational turn-taking may be one of the most well-known interpersonal phenomena that follows this concactenous organization (Levinson, 2006), and the early development of one of its temporal features – the duration of turn-transitions – will be the focus of the two empirical studies in this dissertation.

Interactive Contingency

As used in microanalytical research (Beebe et al., 2016; Jaffe et al., 2001), the concept of interactive contingency is fundamentally related to the predictability in interpersonal behavior. Jaffe et al. (2001) argues that for partners to coordinate in timing they must each predict the temporal pattern of one another. In this way, interactive contingency can be described as the way in which a partner's behavior can be predicted from the behavior of the other interaction partner. Beebe et al. (2016) further distinguishes between self-contingency and interactive contingency, and their mutual role in promoting predictability in interpersonal relation. Self-contingency, the relation of one's behavior to one's own previous behavior, must be a relevant element of interactive contingency in the way it promotes predictability of each partners behavior; additionally, examining self-contingency can also be informative of the directionality in which a partner adapts to another and each moment-to-moment role in interpersonal coordination (Beebe et al., 2016; Burgoon et al., 1995).

These operationalizations of self- and interactive contingency are closely tied to the analytical methods employed to model such temporal relationships, particularly in the analysis of dyadic time series (Beebe et al., 2016). Lagged auto-correlation and cross-correlation are, respectively, an example of statistical measures of self- and interactive contingency, and the modelling of the moment-to-moment dynamics of interpersonal coordination usually account for both to distinguish between spurious and actual adaptation and their directionality (Dean & Dunsmuir, 2016; Moulder et al., 2018).

Methodological Aspects of Interpersonal Coordination

Although it is a core dimension of interpersonal coordination, studying the temporal structure of interactive behavior is however a major challenge (de Barbaro et al., 2013; Fogel, 1977; Xu et al., 2020) due to the properties of interactive behavior and the data it produces.

Properties of Interactional Behavior

Interactional behavior is inherently interpersonal – since it requires at least two participants; and multimodal – since it unravels across different channels of behavior (e.g., voice, gesture, gaze). To consider the temporal structure of interactive behavior one must examine the moment-to-moment changes in behavior of both partners, across channels. Meanwhile, interactional partners can at any moment rearrange the multimodal stream of behaviors, by changing – unilaterally or not – their behavior on one or more channels, in the same or in a different direction (Burgoon et al., 1995; Cappella, 1981). A large range of interpersonal adaptation patterns can be derived from these conditions (Burgoon et al., 1995). Additionally, interactive behavior is also stochastic, exhibiting rhythmic qualities that are irregular or non-periodic (Jaffe et al., 2001). The strength of interactive contingency is not constant and unravels in bursts, punctuated by periods of silence.

This poses multiple difficulties to the understanding of the moment-by-moment dynamics of interpersonal coordination. How should the multimodal stream of behavior be recorded and coded? How should the large range of potential relations between them be explored? How non-stationary multivariate data of this nature should be modeled? Historically, microanalysis was one of the first methodological approaches to dabble with such questions.

Microanalysis

Microanalysis is a methodological approach to the research of interactional behavior, often compared to a social microscope, for its power in detailing the moment-by-moment dynamics of social interaction. Microanalysis started as a technique for the detailed observation and coding of recorded interactional behavior (McQuown, 1971) and quickly became a framework for the empirical study of the mostly nonverbal behavior in mother-infant interactions (Bateson, 1971; Stern, 1971). Historically, it has benefited from an extraordinary effort of multidisciplinary integration (McQuown, 1971), as well as the exploration of the potential of the recording technologies and reproduction techniques available (Bull, 2002; Stern, 2002).

As an approach, microanalysis focus on the study of the whole interaction, as it naturally occurs. Because of that, it privileges the dyadic nature of interactional behavior, the natural context where it unravels, and the detailed analysis of the relationships between multiple streams of behavior (Jaffe et al., 2001; Stern, 2002). And favor the technologies and techniques that increase the resolution of that analysis (Cassotta et al., 1964; Condon, 1970; Condon & Ogston, 1966, 1967).

It has been instrumental to the developmental study of mother-infant interactions, helping to uncover elements of its structural and temporal organization (Brazelton et al., 1974; Condon & Sander, 1974; Fogel, 1977; Stern et al., 1975); and improve our understanding of the effects of mother-infant interpersonal coordination in the infant's cognitive and social-emotional development (Beebe et al., 2010; Beebe & Steele, 2016; Feldman et al., 1996; Feldman & Greenbaum, 1997; Jaffe et al., 2001).

However, despite the advantages of a microanalytic approach to the study of interactional behavior, there has been limitations to its full application: (1) most studies were conducted in laboratory environments that had only some of the properties of natural environments; (2) samples were small and cases studies abundant; (3) the variety of channels explored became narrower; (4) the analysis of the temporal organization was limited. These limitations can be attributed to the limitations in the mobility of recording equipment; the lack of options to automate the taxing process of coding; and the statistical limitations discussed above to the modelling of the complex interactional data (Jaffe et al., 2001).

In most of this dissertation we will focus on only one channel of interactional behavior: vocal behavior; analyze a very specific aspect of vocal coordination: turn-taking; and measured it by a simple metric – turn-transition duration – that inherently gives a reliable measure of temporal coordination (Jaffe et al., 2001). Thus, we can only classify loosely our approach as microanalytic; and mostly because of some of the strategic options in the measurement and coding of our data and the framework we will utilize to interpret others' and our own research. Nevertheless, one chapter of this dissertation will be dedicated to a methodological overview of microanalysis.

Conversational Turn-Taking

We will proceed by focusing on the object of our empirical studies – turn-taking. We will explore the essential properties of turn-taking; the most influential models in turn-taking research; the evolutionary, cross-cultural and developmental foundations of turn-taking; and the developmental trajectories of turn-transition durations.

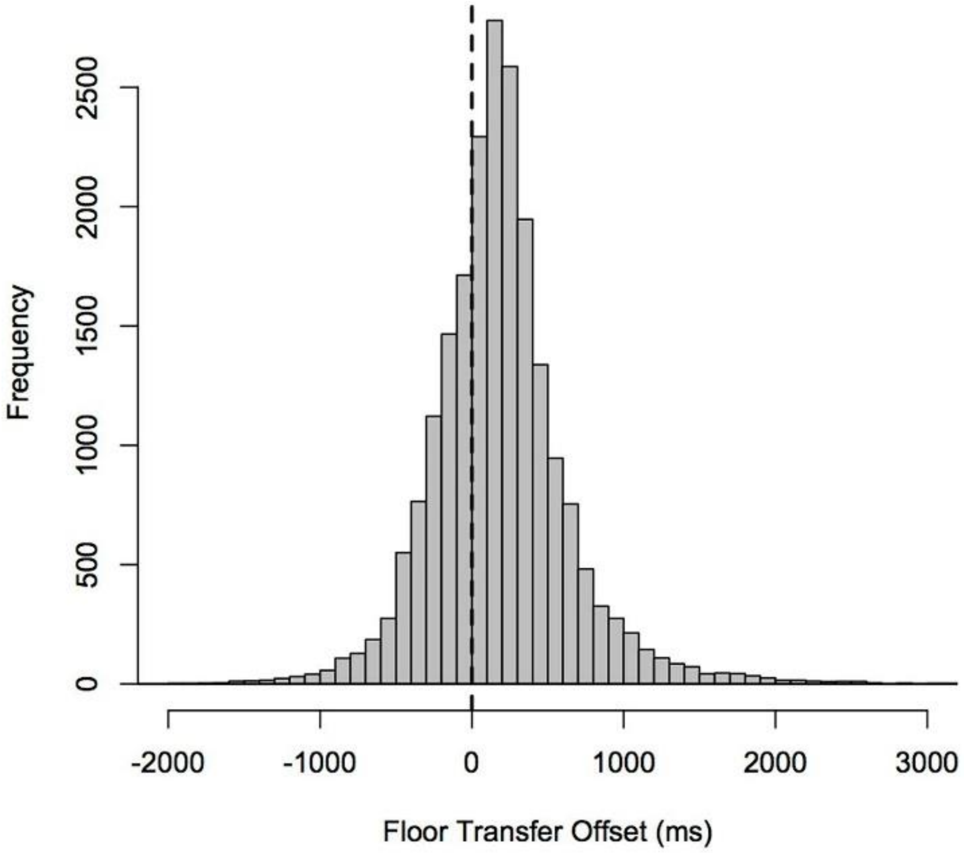
Properties of Turn-taking

The turn-taking organization is characterized by a recurrent pattern of alternation and the minimal-gap minimal-overlap between transitions (Sacks et al., 1974). Turn-transition duration has been used as an effective metric to understand the scale at which the temporal coordination required to the smooth transitions to occur between partners with minimal superposition or silence between turns (Levinson & Torreira, 2015).

By measuring the duration of the silences between turn-transitions – gaps – and the duration of the superpositions that effectively cause a turn change – overlaps –, one can ascertain turn-transition duration. By convention, those measures are standardized by grouping gap and overlap durations in the same temporal scale, a metric known as floor-transfer offset (FTO) – where overlaps assume negative values and gaps assume positive values (Levinson & Torreira, 2015). Gap durations average around 200ms and overlaps around 275ms, with a greater proportion of gaps (60-70%) than overlaps (30-40%); modal FTO has a 0-200ms duration, with a characteristic unimodal and slightly asymmetrical distribution, as illustrated in Figure 1 (Heldner & Edlund, 2010; Levinson & Torreira, 2015).

Figure 1

Floor-Transfer Offsets (FTOs) in the Switchboard Corpus



Note. Histogram of floor-transfer offset (FTO) in the Switchboard Corpus (Godfrey et al., 1992; Calhoun et al., 2010). From "*Timing in turn-taking and its implications for processing models of language*", by S. C. Levinson, F. Torreira, 2015, *Frontiers in Psychology, 6,* 731 (https://doi.org/10.3389/fpsyg.2015.00731). Copyright 2015 by Levinson and Torreira.

Models of Turn-Taking

A few models have been considered in turn-taking research, some more descriptive (Sacks et al., 1974) others more explicative (Levinson, 2006, 2019; Wilson & Wilson, 2005), others even with practical applications for research (Cassotta et al., 1964; Jaffe & Feldstein, 1970). We will proceed by describing the most relevant models in turn-taking literature.

Dialogic Systems Approach

One of the earliest approaches to the study of vocal coordination was the dialogical systems approach (Cassotta et al., 1964; Jaffe & Feldstein, 1970) that began in 1960's with the study of the on-off bursts of dialogue in psychiatric research and the application of automated measures to the segmentation and coding of vocal signals. In the dialogical systems approach, dyads are the main unit of analysis and, therefore, the analysis focus on the temporal relations between dyadic states. The Automatic Vocal Transaction Analyzer (AVTA; Cassotta et al., 1964) employs a binary on-off logic to determine, from the vocal signals, when each partner vocalizes and derive four dyadic states of dialog: 0 = both partners are silent; 1 = partner A vocalizes while partner B is silent; 2 = partner B vocalizes while partner A is silent; 3 = both partners vocalize simultaneously. By using a rule that gives turn ownership to the partner that vocalizes alone until the other partner vocalizes alone instead (Jaffe & Feldstein, 1970), five other dyadic states can be derived: vocalization (V), pause (P), non-interruptive simultaneous speech (NSS), interruptive simultaneous speech (ISS), and switching pause (SP). ISS and SP are the dyadic states that effectively correspond to turn-transitions, respectively, overlaps and gaps (Jaffe et al., 2001).

This model offers several practical advantages to turn-taking research. Since segmentation and coding processes are automated, it accelerates those processes considerably. By relying only on the analysis of vocal signals it is particularly suited to the study of nonverbal vocalizations, specifically in mother-infant interactions or throughout the language acquisition development. Finally, it automatically provides a set of dyadic states, beyond gaps and overlaps, that can be related and provide further information on turn-taking and related phenomena (Beebe et al., 2010; Jaffe et al., 2001).

In the empirical studies of this dissertation, this model will be used to code dyadic states of turn-taking, and extract gap and overlap durations.

Conversation Analysis Approach

The descriptive model developed by Sacks et al. (1974) is considered the standard model of turn-taking (Levinson & Torreira, 2015). It proposes turn allocation rules to accommodate the following set of axioms about turn-taking phenomena: (1) recurrent change in speaker; (2) predominance of one speaker at a time; (3) common, but brief, occurrences of more than one speaker at a time; (4) common transitions with no (or slight) gap or overlap; variable (5) turn order and (6) turn size; absence of specification in advance of the (7) length and (8) content of conversation; and (9) absence of specification in advance of distribution of turns between parties; (10) variable number of parties; (11) talk can be continuous or discontinuous; (12) use of turn-allocation techniques; (13) employment of various “turn-constructional units”; (14) existence of repair mechanisms to deal with errors and violations (Sacks et al., 1974).

As proposed in this model, turn allocation is prescribed by selection rules followed at the end of a turn unit, which is considered a “transition relevant place”: (1) the current speaker may select the next speaker, then the speaker must stop speaking, and the selected listener should start instead; (2) if not selected by the current speaker, any current listener can self-select and gain the new turn by starting to speak; (3) or if none of the current listeners self-select, the current speaker may self-select and continue speaking (Sacks et al., 1974).

The standard model of turn-taking has implications to the language processing occurring in turn-taking (Levinson & Torreira, 2015). The speed of turn-transitions are puzzling, given that for adults it takes at least 600ms to plan and execute the shortest turn (Levinson, 2013), suggesting an overlap between language comprehension and production processes. Specifically, this level of coordination implies that for turn-transition to occur at such speeds, one must be able to predict when another’s turn will end, and simultaneously prepare what to say in the next turn (Levinson & Torreira, 2015).

The implications of this model will be further explored when discussing the interaction engine hypothesis (Levinson, 2006, 2019), below.

Coupled Oscillators Approach

An alternative approach to the temporal coordination dynamics of turn-taking, was proposed by Wilson and Wilson’s (2005) coupled oscillators model, which proposes that interactive partners have internal oscillators that become rhythmically copulated through interaction and negotiation of a shared tempo, in a cyclical phase counter-phase dynamic. For turn-taking, the authors suggest that the syllable production tempo governs the interlocutors readiness to speak, so that the beginning of a syllable from one interlocutor coincides with another's least readiness to speak, and inversely the end of the syllable coincides with increased readiness; by this mechanism simultaneous speech would be minimized (Wilson & Wilson, 2005).

Although there has been support for a coupled oscillator model of turn-taking in primates (Takahashi et al., 2013) and evidence that conversational partners may adjust temporal aspects of their speech (Cappella, 1981; Fusaroli & Tylén, 2016; Manson et al., 2013), there has been no evidence to support a relation between syllable rate and turn-taking (O'Dell et al., 2012).

Interaction Engine Approach

Building on the psycholinguistic implications of the standard model of turn-taking (Levinson & Torreira, 2015; Sacks et al., 1974), Levinson (2006) proposes the interaction engine hypothesis to postulate the existence of universal communicative abilities in humans that precede language in ontogeny and phylogeny, and are the foundations for the predictive requirements of fast turn-transitions. Three systems – turn-taking, repair and sequence organization – appear to be essential dimensions of human conversation organization (Levinson, 2006, 2019). Central to all human interaction is the face-to-face context, the interactional niche where communication and language develops (Levinson, 2019). The fast turn-transition durations characteristic of human turn-taking, require cognitive abilities such as a sensitivity to turn timing and the understanding of other's communicative intentions (Levinson, 2006, 2019).

From the interaction engine hypothesis a number of predictions heave been derived. First, an evolutionary basis for turn-taking; second, the universality of the turn-taking organization across cultures; third, the early development of the cognitive predictive abilities; fourth, the emergence of turn-taking before language acquisition; fifth, the necessity of integration between the foundational interaction system and the linguistic system throughout language acquisition (Hilbrink et al., 2015; Levinson, 2019).

In the following section we will examine evidence of the first four predictions. Later, we will discuss developmental studies on turn-transition duration that explore the implications of the integration between early turn-taking skills and language processing.

Foundations of Turn-Taking

Following some of the predictions derived from the interaction engine hypothesis (Hilbrink et al., 2015; Levinson, 2019) we will examine some of the evolutionary, cross-cultural and developmental evidences of the universal foundations for conversational turn-taking.

Evolutionary Evidence

Evolutionary evidence of some extent of turn-taking has been gathered from the study of insects, amphibians, birds and mammals, with the greatest amount of evidence collected from non-human primates (Pika et al., 2018).

Human's closest phylogenetic relatives, great apes – such as orangutans, bonobos, and chimpanzees – appear to vocalize in response to important events, but not exhibit the flexibility to use these vocalizations voluntarily to communicate (Tomasello & Call, 2019). Meanwhile, their social exchanges are accompanied by instrumental gestures (Halina et al., 2013; Pika et al., 2005) in action-response sequences that alternate at human vocal alternation speed (Rossano, 2013; Rossano & Liebal, 2014). A comparison between bonobos and chimpanzees, showed that these sequences are systematic and varied between species, with bonobos showing more humanlike gaze engagement and speed of response (Fröhlich et al., 2016).

Among other primates, vocal turn-taking has been reported in species of lemurs (Méndez-Cárdenas & Zimmermann, 2009), marmosets (Chow et al., 2015; Snowdon & Cleveland, 1984; Takahashi et al., 2016), titi monkeys (Müller & Anzenberger, 2002), squirrel monkeys (Symmes & Biben, 1988), guenons (Lemasson et al., 2011), and gibbons (Geissmann & Orgeldinger, 2000). Interestingly, this vocal turn-taking appears to be strongly associated with pair-bonding species and subjected to learning processes (Levinson, 2019). As evidenced with the guenon species – *Cercopithecus campbelli* – which albeit having the instinct to reciprocate, must learn the timing of their response and to avoid overlaps (Lemasson et al., 2011). Or the marmoset species – *Callithrix jacchus* - where mothers reinforce learning by penalizing (with nonresponse) infant overlaps by and interrupting the wrong type of responses from their infants (Chow et al., 2015; Takahashi et al., 2016).

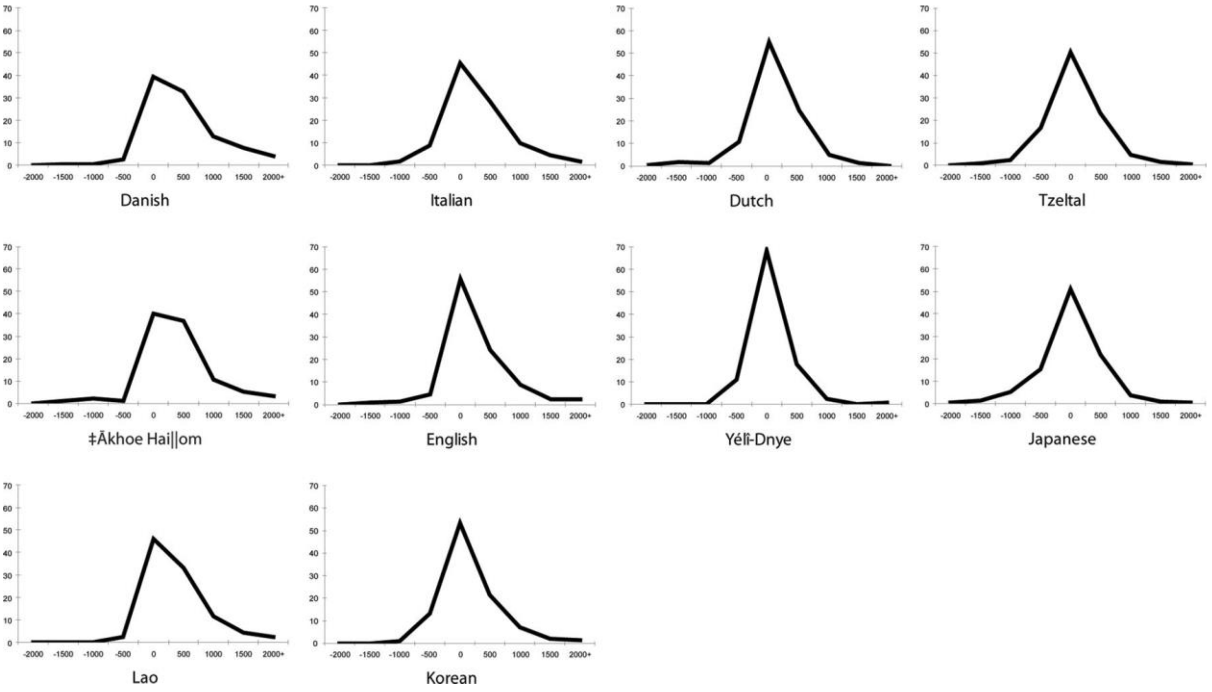
Overall, there is increasing evidence of gestural and vocal turn-taking across primate species pointing to a common ancestry of turn-taking - although its independent evolution within species cannot be ruled out (Levinson, 2019).

Cross-cultural evidence

In a major cross-cultural study examining turn-taking across 10 languages varying in type, geographical location and cultural setting, Stivers et al. (2009) examined previous anthropological claims of substantial differences in turn-taking timing between different cultures. Results shown evidence of the universality of the minimal-gap minimal-overlap pattern characteristic of turn-taking, as well as of the typical slightly unimodal and asymmetrical distribution, with the highest number of turn-transitions occurring between 0 and 200ms (Figure 2). All languages show on average a small positive offset, and although there were differences in average offset, those were within a range of 250 ms from the cross-language mean, and not as substantial as previously suggested. Furthermore, the same factors contributed for speed timing across cultures (Stivers et al., 2009).

Figure 2

Distribution of Turn-Transitions for 10 Languages



Danish
Italian
Dutch
Tzeltal
‡Ākhoe Hai||om
English
Yélî-Dnye
Japanese
Lao
Korean

Note. Distributional plots of turn-transitions in 10 languages, showing similar unimodal [0-200 ms] distributions across languages. From "Universals and cultural variation in turn-taking in conversation", by T. Stivers, N. J. Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, F. Rossano, J. P. Ruitera, K.-E. Yoon, S. C. Levinson, 2009, *Proceedings of the National Academy of Sciences (PNAS)*, *106*(26), 10587-10592 (https://doi.org/10.1073/pnas.0903616106). Copyright 2009 by Stivers, Enfield, Brown, Englert, Hayashi, Heinemann, Hoymann, Rossano, Rutiera, Yoon and Levison.

Developmental evidence

Evidence of the early emergence of the turn-taking was first reported from Bateson's (Bateson, 1971, 1975) seminal microanalytic work on the vocal exchanges from naturalistic mother-infant interaction, known as proto-conversations for their resemblance to the temporal structure of adult conversations. Results suggested as early as 1.5 months, infants vocalize by turns, and that this alternation pattern is already the prevalent mode of organization in proto-conversations. Later studies suggest that mothers may promote this alternation structure by contingently responding to infant's vocalizations and by producing longer pauses after their vocalizations to increase the likelihood of infants engagement in the conversation-like structure (Elias et al., 1986). Other studies of infant's vocal response to social stimulation have indicated that contingent stimulation influences the distribution of infants vocalizations, structuring it towards a burst-pause pattern, from at least 3-months-old (Bloom, 1977; Masataka, 1993). Further microanalytic work, also demonstrated that the interactive contingency in mother-infant interaction can been found across different modalities of communication (Brazelton et al., 1974; Fogel, 1977; Kaye & Fogel, 1980; Stern, 1971; Tronick et al., 1977).

Evidence of infant sensibility to the interruption of temporal contingency in mother-infant interaction was first gathered by Murray and Trevarthen (1985) using an adaptation of the still-face paradigm (Tronick et al., 1978) to examine the effect of the introduction of natural and unnatural breaks of contact in mother-infant interaction in infants' affect and self-regulatory behavior, at 2 and 3-months-old. Results revealed that, when mother's would unnaturally break the interaction by becoming unresponsive (blank-face), infants reduce positive affect and engage in more self-regulatory behaviors than when mothers naturally break contact by turning away to interact with an experimenter. Further, when the temporal relation in mediated interactions is manipulated, infants gaze less at their mothers and present a more detached pattern of affect when compared to the reaction to the blank-face manipulation. Although following research disagreed on the exact age at which this sensitivity to the interruption of contingency in social interaction emerges (Bigelow & DeCoste, 2003; Nadel et al., 1999; Rochat et al., 1998), there is evidence of such sensitivity from at least 3-months-old (Striano et al., 2006).

Evidence of early development of the cognitive ability to predict turn endings was gathered by Casillas and Frank (2013, 2017) by tracking the eye movement of 1 to 7-years-old children, when presenting videos of puppet conversations. Results demonstrated that even 1 and 2-years-old infants are able to accurately and spontaneously predict the occurrence of turn-transitions by gazing at the upcoming speaker before it begins its turn. Similar results can be found when 1 and 2-year-olds observe videos of two people interacting while facing each other (Thorgrimsson, 2014). Further evidence of infant's early predictive ability, can be found in that study. By manipulating the type of vocalization emitted (uttered sentence or non-speech sound) and the response of the recipient (no response), infants shift gaze to the recipient, quicker and for longer durations, when an uttered sentence was spoken. Suggesting that by 12-months-old, infants not only are able to predict turn ending, but also distinguish between communicative intentions.

Next, we will focus on the developmental trajectory of turn-taking.

Developmental Trajectory of Turn-taking

The study of the developmental trajectory of turn-taking has been clouded by some of the limitations of early work: with few longitudinal data, small samples, different metrics and the focus on only gaps (Bateson, 1975; Beebe et al., 1988; Elias et al., 1986; Garvey & Berninger, 1981; Jasnow & Feldstein, 1986) or overlaps (Elias & Broerse, 1996; Ginsburg & Kilbourne, 1988; Rutter & Durkin, 1987). These limitations have produced inconsistent results among gap durations, and an incomplete picture of the role of overlaps to turn-transition durations.

More recently, the developmental trajectory of turn-taking has been recently explored through finer-grained analysis of turn-transition durations, improving on previous research either by measuring gaps and overlaps, providing longitudinal data with several age points, or larger cross-sectional samples (Casillas et al., 2016; Hilbrink et al., 2015; Stivers et al., 2018).

Hilbrink et al. (2015) studied the durations of both gaps and overlaps in mother-infant free-play interactions. Using a longitudinal design, 12 dyads were assessed at six age points (3, 4, 5, 9, 12, 18 months). Median durations were reported for infants and their mothers. Infant gap durations evolved from around 600 ms (3-5 months) to significantly increase to around 1100 ms (9 months) and 975 ms

(12 months), and finally decreased to approximately 700 ms (18 months). Meanwhile, infant overlap durations remained quite stable over time, particularly from 9 months onwards (~600 ms). Children's turn-transitions durations were – with the exception of 5 months gaps – consistently longer than the durations of their mothers. Significant effects were found for age and the interaction between age and person (infant, mother), for both overlap and gap duration.

Casillas et al. (2016) analyzed the response latency in question-response pairs selected from the conversations of 5 children with their caregivers. Six age points were examined from one year and 8 months (1;8) to three years and 5 months (3;5). Response latencies to questions were measured by aggregating gap and overlap durations. When considering all six combinations of question type (yes/no, wh-) and answer complexity (3 levels) analyzed, results show no age effects, with a median response latency of 575 ms across ages. Still, the level of answer complexity had a significant effect on children's response latency, with shorter durations for simpler than complex levels. Meanwhile, when analyzing only the yes/no type of question-answer pairs, researchers have then found a significant effect of age, of response complexity, and of the age by response latency interaction. Children's median response latencies significantly decreased from the first age point (651 ms) to the last age point (469 ms), and more complex responses had significantly longer latencies than simpler responses.

Stivers et al. (2018) studied the response latency in pair-adjacent (question-response) turn-transitions from children triadic conversations instead, in a cross-sectional sample of 95 children from 4 to 8 years old, grouped in same-graded groups of three. For the analysis, children were grouped in two age ranges (4-5 years, 6-8 years). Younger children displayed longer median response latencies (500 ms), than older children (400 ms). Additionally, results shown adult-like (Stivers et al., 2009) distributional differences between types of pair-adjacent transitions, for both age groups. With non-answer responses being significantly longer than answer responses; disconfirmations were significantly longer than confirmations; and interjections significantly faster than other answer types.

A very recent meta-analysis of 26 turn-taking studies with typical and atypical populations analyzed a total of 78 estimates of response latencies (using gaps only) in adult-child interactions, with child ages ranging from 0 to 96 months (Nguyen et al., 2022). The analysis predicted a gradual increase in the trajectory of gap durations, up until 40 months.

Present Work

In this section we will review some of the current challenges in the study of turn-taking's developmental trajectory that we will like to address in this dissertation. A summary of the studies that will be presented as original work of this dissertation, will be offered at the end.

Challenges in Developmental Turn-taking Research

The most recent studies in the developmental trajectory of turn-transition duration (Casillas et al., 2016; Hilbrink et al., 2015; Stivers et al., 2018), follow a markedly psycholinguistic approach and examine some of the language processing predictions of the interaction engine hypothesis (Levinson, 2006, 2019). On one side, by exploring evidence that the integration between infant's foundational interaction system and the development of the language system would produce an increase in turn-transition duration, due to additional cognitive processing (Hilbrink et al., 2015). On the other side, by probing when this integration is completed to a point that turn-transition durations converge to the adult minimal-gap minimal-overlap standard (Casillas et al., 2016; Lindsay et al., 2019; Stivers et al., 2018).

We will analyze this research through a microanalytic lens and present, when possible, arguments for alternative interpretation or the benefits of a different approach.

The Effect of Object-play

In Hilbrink et al. (2015), even though the authors expected an increase in gap duration by 12 months due to predicted additional cognitive effort from language processing, the longer gap durations at 9 months instead, are interpreted as evidence that the integration between the interactional system and the language system starts affecting turn-transition duration at an earlier age.

We find this interpretation problematic. Indeed, it is not unfeasible that some integration between the interaction system and the linguistic system is already occurring by 9 months old. However, the type of linguistic taxing expected by the authors (Hilbrink et al., 2015) would suggest the development of comprehension and production processes that are not yet apparent at 9 months – when infants' communication abilities are mostly nonverbal. Additionally, since there were no measurements of gap and overlap duration in-between 5 and 9 months, there is no way to discard that the reported increase in gap duration at 9 months would not be already apparent from an earlier age point in-between (i.e., at

6, 7 or 8 months). Evidence of an earlier increase in gap duration would be highly improbable to attribute to linguistic processing.

Furthermore, using the lens of the microanalytic approach, a fundamental difference in the context where 3- to 5-months old and 9- to 18-months-old interactions were examined can be spotted. In Hilbrink et al. (2015) appropriate adaptations were made to the setting and materials to account for infant development. Younger infants (3- to 5-months) played face-to-face with their caregivers while sitting supported and stationary. Older infants (9 to 18 months) were free to move and interact with the environment. In both settings, age-appropriate toys were available to play.

Nonetheless, throughout the second half of infant's first year, motor developments progressively capacitate infants to support their body weight, locomote, reach to and manipulate objects, providing increasing ability and agency for infants to engage in the interaction with objects (Adolph & Hoch, 2019; Ruff, 1984). Moreover, during the same period, infants develop the ability to engage in triadic interactions, where infants and caregivers jointly interact with an object (Bertenthal & Boyer, 2015); as, well as a preference for this type of object-play (Bakeman et al., 1990; Williams, 2003).

Given the ability, preference and opportunity to freely engage in object-oriented interactions, we note that the context of interaction where older infants were examined appears to favor object-oriented interaction over face-to-face interaction. We conjecture that the centrality of object play, particularly throughout the second half of the first year when the key developments are occurring, may have an effect in the flow of the interaction, with similar increases in gap duration, to those reported at 9 and 12 months (Hilbrink et al., 2015).

In Study 1 of this dissertation (Chapter II) we will test (1) how gap and overlap durations, at an age before 9 months, compare to their durations, at an age after 9 months; and (2) how object-play may affect turn-transition duration.

The Difference Between Gaps and Overlaps

An essential contribution of Hilbrink's et al. (2015) study, though, was to demonstrate that gaps and overlaps could have a different developmental trajectory. While gap duration appears to have significant increases and decreases throughout infancy, overlap duration maintains quite stable over time. Another overlooked aspect of this difference is that the proportion of gaps and overlaps in infant's turn-transition appears already to be very adult-like, with overlaps occurring only 20-40% of the times.

Even though this relevant distinction was made in infancy, psycholinguistic research, after 18 months, do not highlight the differences between the trajectories of gaps and overlaps. Instead, response latency – is used to measure turn-transition duration, which like the floor-transfer offset (FTO) metric aggregates gap and overlap durations in the same temporal scale – with the former assuming positive values, and the latter assuming negative values (Casillas et al., 2016; Stivers et al., 2018). If, in one hand, this facilitates comparisons with adult FTO durations; on the other hand, the differences in the developmental trajectories of gaps and overlaps is obscured, if durations are not reported separately for gaps and overlaps.

Behind this choice are two assumptions. First, that overlaps are a measure of anticipation in the cognitive processing, and should be measured as negative gaps. Second, there is also the assumption that enough evidence has been provided in Hilbrink et al. (2015) that overlap durations converge to adult standards earlier in development – despite infant's median overlap duration (~575ms), from 9 to 18 months, still being longer than the 205ms median in adults (Levinson & Torreira, 2015).

In Casillas' et al. (2016) longitudinal study, with six age points between 20 and 41 months, the authors probe for evidence that toddlers are improving their turn-transition timing and converging to the minimal-gap minimal-overlap pattern of adult turn-taking. No age effect was found when considering all types of turn-transitions considered, with a median response latency of 625ms, across age points.

If we take into account (1) the timings of 18-month-old infants gaps (~700ms) and overlaps (~575ms) reported in Hilbrink et al, (2015); and (2) the assumption that overlap duration is already converging to adult standard – and by that, not getting longer or more frequent. The median response latency - aggregating gaps and overlaps – of 625 ms, reported in Casillas et al. (2016), would perhaps suggest a relevant increase in gap duration in the short interval between both studies. This increase would be more in line with the taxing effect of linguistic processing that was expected earlier in infants, and assumed to be evident by 9 months (Hilbrink et al., 2015). It would also be consistent with the latest meta-analysis that predicts an increasing trajectory in gap duration, at least until 40 months (Nguyen et al., 2022). Nonetheless, this possibility is unaccounted in the work of Casillas et al. (2016), which focus instead on evidence of an improvement in children's turn timing.

Similarly, Stivers et al. (2018), when studying older children (4-8 years), only present aggregated results of response latency, showing only a slight decreasing with age, with minimal age differences between the 4-5 years-old group (500ms) and the 6-8 years-old group (400ms).

We conjecture that if gap and overlap would be analyzed separately in children as in infancy, a better understanding of the contribution of each dimension to turn-transition timing would be possible. Moreover, by using similar metrics it would facilitate comparisons between studies and highlight continuities in the developmental trajectory of turn-transition durations.

In Study 2 of this dissertation (Chapter III) we will explore (1) the informative value of measuring gaps and overlaps separated or aggregated; and (2) the independent developmental trajectories of gaps and overlaps in childhood.

The Focus on Temporal or Semantic Contingency

A definitive difference between psycholinguistic research in infancy and childhood is the move from nonverbal vocal behavior to the linguistic details of verbal transitions. This goes beyond the unmistakable development of children's linguistic abilities it represents a theoretically-driven shift from temporal contingency to the temporal contingency within semantically contingent transitions (Casillas, 2014). In other words, the shift from temporal coordination in the whole interaction to the coordination in semantically relevant transitions. This leads to an increase in granularity in a narrow slice of the interaction, which has at least produced some interesting results to the resolution of the developmental trajectory of turn-taking.

In Casillas et al. (2016), six combinations of question type (yes/no, wh-) and levels of answer complexity (3 levels) were analyzed. As we have discussed, although an age effect was not found when all combinations were considered, there was an effect of the complexity of the pair-adjacent transitions. Suggesting that response latency duration may be affected by the complexity of answers, with longer latencies for more complex answers. Given the range of linguistic development within the period studied (20-41 months), authors (Casillas et al., 2016) suggest that finer-grained developmental patterns may be obscured by this progression in children's linguistic ability; and that focusing on the least ambiguous question-answer type (yes/no pairs), rather than turn timing as a whole, would increase the detectability of time decreasing patterns.

Further analysis (Casillas et al., 2016), using the least ambiguous question-answer type (yes/no) with two levels of complexity, shown that indeed there appears to be a significant decrease in median response latency between the first (651ms) and the last (469ms) age points; moreover, that answer complexity still has an effect, which also interacts with the age effect, producing longer durations for more complex answers throughout the period.

In Stivers et al. (2018), a greater range of pair-adjacent interactions were studied with older children (4 to 8 years), taking also into account some morpho-syntactic and pragmatic aspects of turn-transitions. Distributional differences were found depending on the type of turn transition – answer responses were faster than non-answer responses; interjections faster than other answer types; and confirmations faster than disconfirmations. These patterns, were consistent with adult distributional patterns regardless of age (Stivers et al., 2009).

The argument is made for a non-linear development in children turn-transition timing (Casillas et al., 2016; Levinson, 2019; Stivers et al., 2018), an hypothesis that has been suggested earlier in literature (Ervin-Tripp, 1979), and has recently received more evidence (Lindsay et al., 2019). This may be an important aspect that only a greater degree of linguistic detail could unveil. However, from a microanalytic perspective, temporal contingency in the whole interaction is relevant and provides a reliable measure of interpersonal coordination (Jaffe et al., 2001). We argue that by focusing exclusively on semantic contingent transitions, psycholinguistic research may be possibly carving away too much of the whole interaction to bring a general understanding of interpersonal coordination. Focusing on a more inclusive spectrum of turn-transitions, instead, may make for a better basis for comparison with the mostly nonverbal turn-transitions in infancy.

To (1) examine the informative value of a more inclusive approach to the study of vocal turn-taking in childhood, and (2) account for the developmental continuities between infancy and childhood, we will use, in Study 1 (Chapter II) and Study 2 (Chapter III), the same microanalytic coding strategy (Cassotta et al., 1964; Jaffe & Feldstein, 1970), where all vocalizations are considered and turn-transitions defined by the interactive vocal behavior of both interactional partners.

The Potential of Microanalysis

We have established previously the restrictions in the approach we will adopt to study turn-taking: we will focus only on vocal exchanges; and interpersonal coordination is explored so far as turn-transition duration provides a reliable measure of temporal coordination between interactional partners.

These restrictions are by no means independent from the methodological challenges we have pointed out to the study of the temporal structure of interactional behavior: (1) recording and coding several streams of behavior; (2) exploring a large range of potential patterns; and (3) the modeling the complexities of interactional data. While classical microanalysis offers a framework for the detailed analysis of interactional behavior, it has also encountered obstacles in the same challenges.

It is out of the scope of this dissertation to solve those obstacles. Nevertheless, in Study 3 (Chapter IV) we will explore some of the methodological options that are becoming available and can help us go beyond vocalizations and turn-transition duration, to understand turn-taking as a multidimensional phenomenon of interpersonal coordination.

Summary of Studies in this Dissertation

We will conclude our introduction to this dissertation, by summarizing the studies that will be presented on the following chapters.

In Study 1: “Turn-taking in Object-oriented and Face-to-face Interactions: A Longitudinal Study at 7 and 12 Months” (Chapter II) of this dissertation we will compare turn-transition durations in object-oriented and face-to-face interactions, in infancy. Using a longitudinal design, with two age points (7, 12 months), we will measure the gap and overlap durations of 25 mother-infant dyads in three tasks: (1) free-play with toys, (2) free-play without toys, and (3) challenging object-play. The effects of task, infant’s age (7, 12 months) and direction of turn-transition (mother to child, child to mother), and potential interactions, will be examined independently for gaps and overlaps (Lourenço, Pereira, et al., 2021).

In Study 2: “Turn-taking in Free-play Interactions: A Cross-sectional Study from 3 to 5 Years” (Chapter III) of this dissertation we will examine the developmental trajectories of gaps and overlaps, in childhood. Using a cross-sectional design, including children from three to five years old, we will examine the vocalizations of 44 mother-child dyads in free-play interactions, where toys are available to play and dyads can engage in conversation. Turn-transition duration will be measured by considering gaps and overlaps both independently, and as an aggregated metric (FTO). The effects of children’s age (3-5 years) and direction of turn-transition (mother to child, child to mother), and potential interactions, will be examined for FTO, gap and overlap duration (Lourenço et al., 2022).

In Study 3: “Advances in Microanalysis: Magnifying the Social Microscope on Mother-Infant Interactions” (Chapter IV) of this dissertation we will present a methodological overview of the microanalytic approach. The historical and theoretical origins of microanalysis as a method for observing and coding interactional behavior will be explored. The transformative impact that this approach brought to developmental science, and particularly to our understanding of the intricacies of mother-infant interaction, will be discussed. As will the methodological challenges that prevented classical microanalytic work to fully overcome the obstacles to the study of interactive behavior. A description of present-day technological innovations and techniques that may improve our ability to capture, explore and model the multidimensionality of interactive behavior is offered. From recent research, two use-cases will be explored, that illustrate the transformative potential of those innovations(Lourenço, Coutinho, et al., 2021).

References

Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, *70*(September 2018), 141–164. https://doi.org/10.1146/annurev-psych-010418-102836

Bakeman, R., Adamson, L. B., Konner, M., & Barr, R. G. (1990). !Kung infancy: The social context of object exploration. *Child Development*, *61*(3), 794. https://doi.org/10.2307/1130964

Bateson, M. C. (1971). The interpersonal context of infant vocalization. *Quarterly Progress Report of the Research Laboratory of Electronics*, *100*, 170–176.

Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, *263*(1), 101–113. https://doi.org/10.1111/j.1749-6632.1975.tb41575.x

Beebe, B., Alson, D., Jaffe, J., Feldstein, S., & Crown, C. (1988). Vocal congruence in mother-infant play. *Journal of Psycholinguistic Research*, *17*(3), 245–259. https://doi.org/10.1007/BF01686358

Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., Bahrick, L., Andrews, H., & Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother–infant interaction. *Attachment & Human Development*, *12*(1–2), 3–141. https://doi.org/10.1080/14616730903338985

Beebe, B., Messinger, D., Bahrick, L. E., Margolis, A., Buck, K. A., & Chen, H. (2016). A systems view of mother–infant face-to-face communication. *Developmental Psychology*, *52*(4), 556–571.

Beebe, B., & Steele, M. (2016). How does microanalysis of mother-infant communication inform maternal sensitivity and infant attachment? *Maternal Sensitivity: Mary Ainsworth's Enduring Influence on Attachment Theory, Research, and Clinical Applications*, *15*(5–6), 141–160. https://doi.org/10.1080/14616734.2013.841050

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. *Fundamentals of Nonverbal Behavior: Studies in Emotion & Social Interaction*, *September*, 401–432.

Bertenthal, B. I., & Boyer, T. (2015). The development of social attention in human infants. In A. Puce & B. I. Bertenthal (Eds.), *The many faces of social attention: Behavioral and neural measures* (pp.

21–65). Springer International Publishing. https://doi.org/10.1007/978-3-319-21368-2

Bigelow, A. E., & DeCoste, C. (2003). Sensitivity to social contingency from mothers and strangers in 2-, 4-, and 6-month-old infants. *Infancy*, *4*(1), 111–140. https://doi.org/10.1207/S15327078IN0401_6

Bloom, K. (1977). Patterning of infant vocal behavior. *Journal of Experimental Child Psychology*, *23*(3), 367–377. https://doi.org/10.1016/0022-0965(77)90032-7

Brazelton, T. B., Koslwski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). Willey-Interscience.

Bull, P. (2002). *Communication under the microscope: The theory and practice of microanalysis*. Routledge.

Burgoon, J. K., Stern, L. A., & Dillman, L. (1995). *Interpersonal adaptation: Dyadic interaction partners*. Cambridge University Press.

Cappella, J. N. (1981). Mutual influence in expressive behavior: Adult-adult and infant-adult dyadic interaction. *Psychological Bulletin*, *89*(1), 101–132. https://doi.org/10.1037/0033-2909.89.1.101

Casillas, M. (2014). Turn-taking. In D. Matthews (Ed.), *Pragmatic development in first language acquistion* (pp. 53–70). Benjamins. https://doi.org/10.1075/tilar.10.04cas

Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn-taking, timing, and planning in early language acquisition. *Journal of Child Language*, *43*(6), 1310–1337. https://doi.org/10.1017/S0305000915000689

Casillas, M., & Frank, M. (2013). The development of predictive processes in children's discourse understanding. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 299–304. http://langcog.stanford.edu/papers/CF-cogsci2013.pdf

Casillas, M., & Frank, M. C. (2017). The development of children's ability to track and predict turn structure in conversation. *Journal of Memory and Language*, *92*, 234–253. https://doi.org/10.1016/j.jml.2016.06.013

Cassotta, L., Feldstein, S., & Jaffe, J. (1964). AVTA: A device for automatic vocal transiction analysis.

Journal of the Experimental Analysis of Behavior, *7*(1), 99–104. https://doi.org/10.1901/jeab.1964.7-99

Chow, C. P., Mitchell, J. F., & Miller, C. T. (2015). Vocal turn-taking in a non-human primate is learned during ontogeny. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1807). https://doi.org/10.1098/rspb.2015.0069

Condon, W. S. (1970). Method of micro-analysis of sound films of behavior. *Behavior Research Methods & Instrumentation*, *2*(2), 51–54. https://doi.org/10.3758/BF03210994

Condon, W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior. *The Journal of Nervous and Mental Disease*, *143*(4), 338–347. https://doi.org/10.1097/00005053-196610000-00005

Condon, W. S., & Ogston, W. D. (1967). A segmentation of behavior. *Journal of Psychiatric Research*, *5*(3), 221–235. https://doi.org/10.1016/0022-3956(67)90004-0

Condon, W. S., & Sander, L. W. (1974). Synchrony demonstrated between movements of the neonate and adult speech. *Child Development*, *45*(2), 456. https://doi.org/10.2307/1127968

de Barbaro, K., Johnson, C. M., Forster, D., & Deák, G. O. (2013). Methodological considerations for investigating the microdynamics of social interaction development. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 258–270. https://doi.org/10.1109/TAMD.2013.2276611

Dean, R. T., & Dunsmuir, W. T. M. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, *48*(2), 783–802. https://doi.org/10.3758/s13428-015-0611-2

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, *3*(3), 349–365. https://doi.org/10.1109/T-AFFC.2012.12

Elias, G., & Broerse, J. (1996). Developmental changes in the incidence and likelihood of simultaneous talk during the first two years: A question of function. *Journal of Child Language*, *23*(1), 201–217. https://doi.org/10.1017/s0305000900010151

Elias, G., Hayes, A., & Broerse, J. (1986). Maternal control of co-vocalization and inter-speaker silences in mother-ineant vocal engagements. *Journal of Child Psychology and Psychiatry*, *27*(3), 409–415. https://doi.org/10.1111/j.1469-7610.1986.tb01842.x

Ervin-Tripp, S. (1979). Children's verbal turn-taking. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 391–414). Academic Press.

Feldman, R., & Greenbaum, C. W. (1997). Affect regulation and synchrony in mother-infant play as precursors to the development of symbolic competence. *Infant Mental Health Journal*, *18*(1), 4–23. https://doi.org/10.1002/(SICI)1097-0355(199721)18:1<4::AID-IMHJ2>3.0.CO;2-R

Feldman, R., Greenbaum, C. W., Yirmiya, N., & Mayes, L. C. (1996). Relations between cyclicity and regulation in mother-infant interaction at 3 and 9 months and cognition at 2 years. *Journal of Applied Developmental Psychology*, *17*(3), 347–365. https://doi.org/10.1016/S0193-3973(96)90031-3

Fogel, A. (1977). Temporal organization in mother-infant face-to-face interaction. *Studies in Mother-Infant Interaction*, *January 1977*, 119–152.

Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, *6*(April), 1–14. https://doi.org/10.1038/srep25887

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, *40*(1), 145–171. https://doi.org/10.1111/cogs.12251

Garvey, C., & Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse Processes*, *4*(1), 27–57. https://doi.org/10.1080/01638538109544505

Gatewood, J. B., & Rosenwein, R. (1981). Interactional synchrony: Genuine or spurious? A critique of recent research. *Journal of Nonverbal Behavior*, *6*(1), 12–29. https://doi.org/10.1007/BF00987933

Geissmann, T., & Orgeldinger, M. (2000). The relationship between duet songs and pair bonds in siamangs, Hylobates syndactylus. *Animal Behaviour*, *60*(6), 805–809. https://doi.org/10.1006/anbe.2000.1540

Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, *15*, 221–235.

https://doi.org/10.1017/S0305000900012344

Halina, M., Rossano, F., & Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Animal Cognition*, *16*(4), 653–666. https://doi.org/10.1007/s10071-013-0601-7

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568. https://doi.org/10.1016/j.wocn.2010.08.002

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology*, *6*(September), 1–12. https://doi.org/10.3389/fpsyg.2015.01492

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), i–viii, 1–132. http://www.ncbi.nlm.nih.gov/pubmed/11428150

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

Jasnow, M. D., & Feldstein, S. (1986). Adult-like temporal characteristics of mother-infant vocal interactions. *Child Development*, *57*(3), 754–761. http://www.jstor.org/stable/1130352

Kaye, K., & Fogel, A. (1980). The temporal structure of face-to-face communication between mothers and infants. *Developmental Psychology*, *16*(5), 454–464. https://doi.org/10.1037/0012-1649.16.5.454

Kendon, A. (1970). Movement coordination in social interaction: Some examples described. *Acta Psychologica*, *32*(C), 101–125. https://doi.org/10.1016/0001-6918(70)90094-6

Lemasson, A., Glas, L., Barbu, S., Lacroix, A., Guilloux, M., Remeuf, K., & Koda, H. (2011). Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Scientific Reports*, *1*, 12–15. https://doi.org/10.1038/srep00022

Levinson, S. C. (2006). On the human interactional engine. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality. Culture, Cognition and Interaction* (pp. 39–69). Berg.

Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 103–130). Willey-Blackwell.

Levinson, S. C. (2019). Interactional foundations of language: The interaction engine hypothesis. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 189–200). MIT Press.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*(JUN), 1–17. https://doi.org/10.3389/fpsyg.2015.00731

Lewkowicz, D. J. (1989). The role of temporal factors in infant behavior and development. *Advances in Psychology*, *59*(C), 9–62. https://doi.org/10.1016/S0166-4115(08)61038-9

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychological Science*, *30*(4), 504–515. https://doi.org/10.1177/0956797618822802

Lourenço, V., Coutinho, J., & Pereira, A. F. (2021). Advances in microanalysis: Magnifying the social microscope on mother-infant interactions. *Infant Behavior and Development*, *64*(April), 101571. https://doi.org/10.1016/j.infbeh.2021.101571

Lourenço, V., Pereira, A. F., Sampaio, A., & Coutinho, J. (2021). Turn-taking in object-oriented and face-to-face interactions: A longitudinal study at 7 and 12 months. *Psychology & Neuroscience*. Advance online publication. https://doi.org/10.1037/pne0000276

Lourenço, V., Serra, J., Coutinho, J., & Pereira, A. F. (2022). *Turn-taking in free-play interactions: A Cross-sectional study from 3 to 5 years*. [Manuscript submitted for publication]. School of Psychology.

Manson, J. H., Bryant, G. A., Gervais, M. M., & Kline, M. A. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behavior*, *34*(6), 419–426. https://doi.org/10.1016/j.evolhumbehav.2013.08.001

Masataka, N. (1993). Effects of contingent and noncontingent maternal stimulation on the vocal behaviour of three- to four-month-old Japanese infants. *Journal of Child Language*, *20*(2), 303–312. https://doi.org/10.1017/S0305000900008291

McDowall, J. J. (1978a). Interactional synchrony: A reappraisal. *Journal of Personality and Social Psychology*, *36*(9), 963–975. https://doi.org/10.1037/0022-3514.36.9.963

McDowall, J. J. (1978b). Microanalysis of filmed movement: The reliability of boundary detection by observers. *Environmental Psychology and Nonverbal Behavior*, *3*(2), 77–88. https://doi.org/10.1007/BF01135606

McQuown, N. A. (1971). *The Natural history of an interview*. University of Chicago Library.

Méndez-Cárdenas, M. G., & Zimmermann, E. (2009). Duetting - A mechanism to strengthen pair bonds in a dispersed pair-living primate (Lepilemur edwardsi)? *American Journal of Physical Anthropology*, *139*(4), 523–532. https://doi.org/10.1002/ajpa.21017

Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychological Methods*, *23*(4), 757–773. https://doi.org/10.1037/met0000172

Müller, A. E., & Anzenberger, G. (2002). Duetting in the titi monkey Callicebus cupreus: Structure, pair specificity and development of duets. *Folia Primatologica*, *73*(2–3), 104–115. https://doi.org/10.1159/000064788

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social Perception in Infants* (pp. 177–197). Ablex Publishers.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science*, *2*(2), 164–173. https://doi.org/10.1111/1467-7687.00065

Nguyen, V., Versyp, O., Cox, C., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the development of turn taking in adult–child vocal interactions. *Child Development*, 1–20. https://doi.org/10.1111/cdev.13754

O'Dell, M., Nieminen, T., & Lennes, M. (2012). Modeling turn-taking rhythms with oscillators. *Linguistica Uralica*, *48*(3), 218–227. https://doi.org/10.3176/lu.2012.3.08

Paxton, A. (2015). *Coordination: Theoretical, methodological, and experimental perspectives.* [Doctoral dissertation, UC Merced]. https://escholarship.org/uc/item/5tx5s7zh

Pika, S., Liebal, K., Call, J., & Tomasello, M. (2005). Gestural communication of apes. *Gesture*, *5*(1–2), 41–56. https://doi.org/10.1075/gest.5.1.05pik

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: Bridging the gap between human and animal communication. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880). https://doi.org/10.1098/rspb.2018.0598

Rochat, P., Neisser, U., & Marian, V. (1998). Are young infants sensitive to interpersonal contingency? *Infant Behavior and Development*, *21*(2), 355–366. https://doi.org/10.1016/S0163-

6383(98)90012-1

Rossano, F. (2013). Sequence organization and timing of bonobo mother-infant interactions. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, *14*(2), 160–189. https://doi.org/10.1075/is.14.2.02ros

Rossano, F., & Liebal, K. (2014). *"Requests" and "offers" in orangutans and human infants*. August*, 335–364. https://doi.org/10.1075/slsi.26.13ros

Ruff, H. A. (1984). Infants' manipulative exploration of objects: Effects of age and object characteristics. *Developmental Psychology*, *20*(1), 9–20. https://doi.org/10.1037/0012-1649.20.1.9

Rutter, D. R., & Durkin, K. (1987). Turn-taking in mother-infant interaction: An examination of vocalizations and gaze. *Developmental Psychology*, *23*(1), 54–61. https://doi.org/10.1037//0012-1649.23.1.54

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Snowdon, C. T., & Cleveland, J. (1984). "Conversations" among pygmy marmosets. *American Journal of Primatology*, *7*(1), 15–20. https://doi.org/10.1002/ajp.1350070104

Stern, D. N. (1971). A Micro-Analysis of Mother-Infant Interaction: Behavior Regulating Social Contact Between a Mother and her 3 1/2-Month-Old Twins. *Journal of the American Academy of Child Psychiatry*, *10*(3), 501–517. https://doi.org/10.1016/S0002-7138(09)61752-0

Stern, D. N. (2002). *The First Relationship: Infant and Mother* (2nd ed.). Harvard University Press.

Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: two modes of communication within the mother-infant dyad. *Annals of the New York Academy of Sciences*, *263*(1), 89–100. https://doi.org/10.1111/j.1749-6632.1975.tb41574.x

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stivers, T., Sidnell, J., & Bergen, C. (2018). Children's responses to questions in peer interaction: A window into the ontogenesis of interactional competence. *Journal of Pragmatics*, *124*, 14–30.

https://doi.org/10.1016/j.pragma.2017.11.013

Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies*, *7*(2), 251–271. https://doi.org/10.1075/is.7.2.08str

Symmes, D., & Biben, M. (1988). Conversational vocal exchanges in squirrel monkeys. In D. Todt, P. Goedeking, & D. Symmes (Eds.), *Primate Vocal Communication* (pp. 123–132). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-73769-5_8

Takahashi, D. Y., Fenley, A. R., & Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693). https://doi.org/10.1098/rstb.2015.0370

Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. *Current Biology*, *23*(21), 2162–2168. https://doi.org/10.1016/j.cub.2013.09.005

Thorgrimsson, G. B. (2014). *Infants' understanding of communication as participants and observers*. [Unpublished doctoral dissertation]. Radbound University Nijmegen.

Tomasello, M., & Call, J. (2019). Thirty years of great ape gestures. *Animal Cognition*, *22*(4), 461–469. https://doi.org/10.1007/s10071-018-1167-1

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Tronick, E., Als, H., & Brazelton, T. B. (1977). Mutuality in mother-infant interaction. *Journal of Communication*, *27*(2), 74–79. https://doi.org/10.1111/j.1460-2466.1977.tb01829.x

Vicaria, I. M., & Dickens, L. (2016). Meta-analyses of the intra- and interpersonal outcomes of interpersonal coordination. *Journal of Nonverbal Behavior*, *40*(4), 335–361. https://doi.org/10.1007/s10919-016-0238-8

Williams, E. (2003). A comparative review of early forms of object-directed play and parent-infant play in typical infants and young children with autism. *Autism*, *7*(4), 361–377. https://doi.org/10.1177/1362361303007004003

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic*

Bulletin and Review, *12*(6), 957–968. https://doi.org/10.3758/BF03206432

Xu, T. L., de Barbaro, K., Abney, D. H., & Cox, R. F. A. (2020). Finding structure in time: visualizing and analyzing behavioral time series. *Frontiers in Psychology*, *11*, 1–22. https://doi.org/10.3389/fpsyg.2020.01457

CHAPTER II – STUDY 1[1]

---

[1] A version of this chapter was published online first in ahead of print, in the American Psychological Association peer-reviewed journal *Psychology & Neuroscience*:

Lourenço, V., Pereira, A. F., Sampaio, A., & Coutinho, J. (2021). Turn-taking in object-oriented and face-to-face interactions: A longitudinal study at 7 and 12 months. *Psychology & Neuroscience*. Advance online publication. https://doi.org/10.1037/pne0000276

Abstract

Objective: The goal of this study was to compare turn-transition duration, for gaps and overlaps, in free-play, with and without objects, in mother-infant interactions, at 7 and 12 months. Few studies have approached the developmental study of turn-transition durations in infancy, and most examine turn-taking only in the context of face-to-face or object-oriented playful interaction. But none has compared the effects of each mode of play in turn-transition duration. ***Methods:*** We analyzed the vocalizations of 25 mother-infant dyads in a semi-structured interaction with three tasks: (1) free-play with toys, (2) free-play without toys, and (3) challenging object play. Gap and overlap frequency and duration were measured, and the effects of age point (7, 12 months), task (1, 2, 3) and direction of turn-transition (infant, mother) assessed using generalized linear mixed modelling for each dependent variable (DV: gaps, overlaps). ***Results:*** Gap durations were substantially shorter, and overlap durations had greater variability, in the free-play without objects condition. There were significant differences between face-to-face and object-oriented conditions for both DV's. Only for gaps, all fixed-effects (age, task, direction), and their two-way interactions, were significant. Post-hoc pairwise comparisons found that, in object-oriented conditions, infants maintain similar durations between age points, unlike their mothers. In the face-to-face condition, infants and mothers had the same descending tendency between age points. ***Conclusions:*** Interactions without objects produce shorter gap durations and may reveal developmental changes in turn-taking. Research should account for the methodological and theoretical implications of this effect.

Keywords: turn-taking, mother-infant interaction, object-oriented, face-to-face, turn-transition

Turn-taking in Object-oriented and Face-to-face Interaction: A Longitudinal Study at 7 and 12 Months

Mother-infant interaction is often compared to an intricate dance across multiple streams of behavior (Stern, 2002). There is a sense of temporal relation between interactive partners, a rhythmicity in their behaviors, where one partner's behavior quickly and contingently follows the behavior of the other (Jaffe et al., 2001). In the speech domain, the rhythmic and sequential nature of mother-infant exchanges is often compared to the organization of adult conversations, and has been described as a proto-conversation (Bateson, 1975; Bruner, 1975) – a pattern of burst and pause in vocalizations, along with partners that engage in turn-taking.

Turn-taking in conversation refers to the recurrent alternation between conversational partners, where one person holds the floor and typically vocalizes alone, followed by a smooth turn-taking transition where partners exchange roles. The turn-transition is tightly coordinated and is characterized by slight or no gaps or overlaps. Gaps are short periods of silence and the most frequent turn-transition; overlaps occur when during the turn-transition, the two partners vocalize simultaneously. The frequency of overlaps is not irrelevant, compared to gaps, but overlaps tend to be very brief in adult-adult conversations (Levinson & Torreira, 2015; Sacks et al., 1974). The precision in the turn-transition is striking. By one measure of turn-taking, the floor-transfer offset, adults in conversation exchange turns with a gap that is on average 200 ms in duration (Levinson & Torreira, 2015). The floor-transfer offset groups overlaps and gaps in the same time scale by representing an overlap duration as a negative value and a gap as a positive value. This measure assumes that adults work towards sustaining a mean turn-transition duration close to zero, in a pattern that has been summarized as “minimal-gap minimal-overlap” (Stivers et al., 2009). The floor-transfer offset has a unimodal and slightly asymmetrical distribution with more positive than negative values (Levinson & Torreira, 2015). Cross-cultural research has shown that this pattern is prevalent across cultures, with the same slightly asymmetrical and unimodal distribution, and a value for the mode of the distribution between 0-200 ms (Stivers et al., 2009). These fast transitions are puzzling, given that it takes at least 600 ms to plan and execute the shortest turn (Levinson, 2013). In order to achieve such smooth turn-transitions, language comprehension and production processes are likely active in parallel, and turn-taking requires the ability of predicting the timing of the turn end and the unfinished information before it ends (Levinson & Torreira, 2015). This sensitivity to turn-timing, and the ability to predict the other’s communicative intentions – the basis for the turn-taking conversation system – can be found already in pre-verbal infants, and may have an evolutionary basis that predates the emergence of language, a conjecture put forward in the interaction engine hypothesis (Levinson, 2006). This hypothesis postulates the existence of universal communicative abilities in humans that precede language in ontogeny and phylogeny (e.g. alternation in vocalization and gesture). These abilities are the cognitive and ethological foundations of conversational turn-taking, and support the interactional niche where language development happens (see Levinson (2019) for a recent review).

Developmentally, characteristics of gaps and overlaps (e.g. proportion during an interaction, mean and median duration) do not necessarily change in the same direction, e.g. gaps increase in mean gap duration without a change in mean overlap duration. For this reason, developmental studies do not use the floor-transfer offset measure, and model gaps and overlaps separately (Hilbrink et al., 2015), or only examine one of the two (Bateson, 1975; Beebe et al., 1988; Elias et al., 1986; Jasnow & Feldstein, 1986).

The turn-taking pattern has been measured in observational studies of adult-infant and adult-child interaction, in free-flow play situations or semi-structured social interaction tasks where experimenters impose some constraints, by providing or removing objects and asking the adult to engage in a particular task. The sensitivity to turn-timing and communicative intentions has also been measured using experimental paradigms. We review both approaches.

Experimental Studies of Sensitivity to Turn-timing and Interactive Contingency

In a study of 2 and 3-month-olds, Murray and Trevarthen (1985) manipulated mother-infant interactions and measured infants' direction of attention, communicative effort, and affect. In a first experiment, they used an adaptation of the still-face paradigm (Tronick et al., 1978) to introduce natural and unnatural breaks of contact into mother-infant interaction. In the blank-faced condition, mothers were asked to unnaturally break the interaction with their infants by posing a still face while still looking at their infants. In the interruption condition, mothers would break contact by turning away to interact with an experimenter that enters the room. Results revealed that, in the interruption condition, although there is a slight decrease of positive affect, distress signals are not present and a reduction in communicative effort is accompanied by a shift in attention towards the experimenter; a different pattern was detected in the blank-face condition, where infants exhibit an increase in communicative efforts, distress signals, self-regulatory behavior and, eventually, withdrawal from the interaction. In a second experiment, Murray and Trevarthen (1985), used delayed recordings to also disrupt the temporal relationships between mother and infant behaviors. In a closed video circuit, mothers would first interact online with their infants; after, a recording of mother's interactive behavior would be replayed to the infant to create an unresponsive interaction. Infants gazed less at their mothers and presented a more detached pattern of affect when compared to the still-face manipulation. The paradigm of Murray and Trevarthen (1985) instigated an entire line of research, and although there are some controversial results (Bigelow & DeCoste, 2003; Nadel et al., 1999; Rochat et al., 1998) there is evidence of infants' sensitivity to the interruption of contingency in social interactions from at least 3-months-old (Striano et al., 2006). Using a different experimental paradigm, Casillas and Frank (Casillas & Frank, 2013) analyzed predictive processing by tracking eye movements while participants viewed videos of puppet conversations. They tested infants and children aged 1 to 7 years of age. This study demonstrated that even 1 and 2-year-old infants are able to accurately and spontaneously predict the occurrence of turn-transitions by gazing at the upcoming speaker before it begins its turn. Similar results can be found when 1 and 2-year-olds observe videos of two people interacting while facing each other (Thorgrimsson, 2014). By comparing when one person uttered a sentence vs. a non-speech sound, while controlling for no response of the other person, Thorgrimsson (2014) was able to demonstrate that infants shift gaze quicker and for longer durations to the recipient of the uttered sentence than when the person vocalizing emitted a non-speech sound, i.e. infants not only are able to predict turn-transitions before they occur, but also distinguish from speech and non-speech vocalizations.

Observational Studies of Turn-taking in Infants

Turn-taking research with pre-verbal infants usually observe vocal behavior in the context of naturalistic or semi-structured playful interactions (Bateson, 1975). With younger infants, this translates to a setup where mothers and infants are seated facing each other and interact face-to-face through different streams of behavior (Beebe et al., 1988). With older infants though, this may translate to a setup were one or more toys are available and these become part of more object-oriented or object-mediated interactions (Jasnow & Feldstein, 1986). These adjustments reflect the development, throughout the second half of the first year, of not only infant's gross motor abilities, that progressively allow infants to support their own body, sit, stand, and eventually walk, but also the finer motor abilities that evolve from simple grasping to object manipulation and inspection, which allows precise active object exploration (Adolph & Hoch, 2019; Ruff, 1984). Infants ability to engage in triadic interactions, involving infants, an object, and another partner, also develops throughout this period (Bertenthal & Boyer, 2015), which in turn translates into a major change in infant's social play patterns, with an increase in the amount of time spent playing with toys in social interactions (Bakeman et al., 1990) and a preference for object-oriented interactions (Williams, 2003), from 6 to 12 months.

Microanalytic studies of mother-infant interactions (see Lourenço et al., 2021 for a review) have provided evidence for infants' sensitivity to contingency in communication. Mother-infant interaction unravels through different streams of behavior and interactive contingency can be found across different modalities of communication (Brazelton et al., 1974; Condon & Sander, 1974; Fogel, 1977; Kaye & Fogel, 1980; Tronick et al., 1977). Studies of infant's vocal responses to social stimulation have shown that, at least by 3-months of age, contingent stimulation modulates the distribution of infants vocalizations, structuring them towards a burst-pause pattern (Bloom, 1977; Masataka, 1993). Bateson's (1971, 1975) seminal work on the temporal structure of mother-infant proto-conversations suggested that, as early as 1.5 months, infants vocalize by turns, and that this alternation pattern is already the prevalent mode of organization in proto-conversations. Further studies suggest that mothers may promote this alternation structure by contingently responding to infant's vocalizations and by producing longer pauses after their vocalizations in order to increase the likelihood of infants engagement in the conversation-like structure of the social exchange (Elias et al., 1986).

Although the fine-grained timing of turn-transitions has been studied extensively in adults (Levinson & Torreira, 2015), equivalent developmental research is still sparse. Studies have examined only one type of turn-transition (gap or overlap), and use different metrics (proportion or duration); sample size is typically small and there are only a few longitudinal studies. This has produced an incomplete picture of turn-taking's developmental trajectory (Hilbrink et al., 2015).

Regarding overlaps, work that measured overlaps usually focus on proportion, rather than duration. The main finding is a developmental curve where overlaps are more prevalent in the first 3 months of life (Ginsburg & Kilbourne, 1988), decrease throughout the first 18 months and then increase (Elias & Broerse, 1996) or perhaps increase even earlier, between 9 and 24 months (Rutter & Durkin, 1987). Regarding the developmental trajectory of gap duration, results are also unclear. In Bateson's (1975) study, only one dyad was assessed at two age points (1.5 and 3.5 months) and infant gap durations was measured by including the timing of mother's previous utterance – the gap duration averaged 1300 ms. In the longitudinal study of (Elias et al., 1986), with a sample of 6 dyads, measured at 3 and 4 months, gap mean duration had a similar value, 1200 ms. Yet, in a study of 4-month-olds , using a similar recording method, (Beebe, 1988) reports an average infant gap duration of 800 ms, which is similar to the mean value of 875 ms in another cross-sectional study, at 9 months (Jasnow & Feldstein, 1986).

An exception to measuring only gaps or only overlaps is a recent longitudinal study that observed 12 mother-infant dyads, at six age points (3, 4, 5, 9, 12, 18 months) and in a free play task (Hilbrink et al., 2015). The main prediction was that the emergence of productive language would coincide with longer gap durations. Concerning overlaps, this study found that 3 to 5-months-old infants overlap with their mothers in a third of their vocalizations, but by 18 months, infants produce overlaps with a proportion similar to their mothers (20%). Median overlap duration remained stable, around 575 ms, particularly from 9 to 18 months. This suggests that although the turn-taking alternation structure seems to became prevalent from 5 months onward, infants are not yet attempting to reduce the duration of their overlaps, like adults do (Hilbrink et al., 2015). As to the developmental trajectory of gaps, median durations from 3 to 5 months varied around 600 ms and then increased at 9 months, to approximately 1100 ms, followed by a slight decrease over time. At 5 months of age, infants were significantly faster than at 9 and 12 months, but not compared with 18 months. The authors argue that the longer durations in gaps are consistent with their predictions, based on the interaction engine hypothesis (Levinson, 2006), and studies of gap duration in older children (Casillas et al., 2016; Garvey & Berninger, 1981).

The Hilbrink et al. (Hilbrink et al., 2015) study only tested dyads in a free-play with objects task (but took into account the infants' motor ability and the age-appropriateness of toys provided). However, it cannot be used to directly compare object-oriented and face-to-face (non-object oriented) interactions, that is, we cannot compare the effect of object play on turn-transition durations, particularly when there may be developmental differences to infants' engagement in object-oriented interactions throughout the age points assessed.

Present Study

We measured the timing of turn-transition, separately for overlaps and gaps, while infants engaged with their mothers in object-oriented play interactions and play interactions without objects. Infants were observed at two age points; first, at 7 months – when infants are preverbal but engage in intense active object exploration; second, at 12 months – when infants start producing language. We asked mother-infant dyads to engage in three tasks: one was free play with a set of age-appropriate toys, a second was free play without any toys present, and a third was object-play using a toy that was above the infant's developmental level.

Overall, we expect to find evidence, both in terms of frequency and duration, that turn-transitions are different whether object manipulation is present or absent in the interaction. We propose three specific hypotheses for gap durations, and one for overlap durations.

For gap duration, we conjecture that object-oriented interaction poses additional challenges to maintaining short turn-taking gaps. Therefore, we expect (Hypothesis 1) to find longer gap durations in the object-oriented conditions, than in the face-to-face condition, regardless of infant's age. We will be using the findings of the Hilbrink et al. (2015) study as a reference for the object-oriented conditions (although in that study, object manipulation was not controlled for, as mothers had toys available but were free to choose what to do during the interaction). We therefore expect that infant's gap durations, at 7 and 12 months, will be, only for the object-oriented conditions, closer to the median durations found previously in the Hilbrink et al. (2015) study: 9 months (~1100 ms) and 12 months (~975 ms). For the face-to-face condition, we expect that infant's gap durations will be closer to the average values of adult floor-transfer offsets (~200 ms), as reported by Levinson & Torreira (2015), although gap duration in our study will be likely higher, since the floor-transfer offset measure includes gaps and overlaps.

Regarding the developmental trajectory of gap durations, we expect a different trajectory for object-oriented and face-to-face conditions. We conjecture (Hypothesis 2) that, for object-oriented conditions, gap durations will be similar at 7 and 12 months. In the only comparable study, Hilbrink et al. (2015), there was a significant increase in gap duration, between 5 to 9 months, and no significant differences between 9 and 12 months. Around 7 months, there are major developmental transitions in motor development (Adolph & Hoch, 2019; Ruff, 1984) that support a greater ability to engage in triadic interactions (Bertenthal & Boyer, 2015); this is confirmed by an increased preference for object-oriented play (Bakeman et al., 1990; Williams, 2003). Therefore, we conjecture that, only for the object-oriented conditions, gap durations should be similarly affected by the presence of object manipulation, at 7 or 12 months. On the other hand (Hypothesis 3), for the face-to-face condition we conjecture that there is a developmental decrease in gap duration when object manipulation is absent. We speculate that the development of infant's communicative abilities from 7 to 12 months supports the fluidity of the interaction resulting in shorter gap durations. The ability to understand the communicative intentions of their mothers, to engage in joint attention and to point (Bakeman & Adamson, 1984; Carpenter et al., 1998), are still mostly pre-verbal developments that should not tax gap durations with the same demands of language processing reported in studies with older children (Casillas et al., 2016; Garvey & Berninger, 1981).

Finally, in terms of overlaps, and given the limitations of previous research – often circumscribed to studies that only measured frequency – we can only conjecture based again on the results found in Hilbrink et al. (2015), where no significant differences were found for infant overlap durations. We have no reason to expect (Hypothesis 4) any significant difference in infants' overlap duration, at least for object-oriented interactions. If we take the value reported in Hilbrink, et al. (2015) at 12 months, as the reference value for overlap duration, the median duration is approximately 575 ms.

Method

Participants

The data reported here is part of a larger longitudinal research project studying affective touch processing that evaluated infants at three different stages: 7, 12, and 18 months (Miguel et al., 2019, 2020; Serra et al., 2020). Mothers and their infants were initially recruited in parenting classes, social networks, and daycare centers in Braga, Portugal. In the present study we examined a random subset (n = 25) of the 41 dyads that completed all tasks at both 7 and 12 months old, and exhibited at least one turn-transition by condition. All infants were typically developing infants with normal birth weight (> 2500 g; two infants had slightly lower birth weights: 2350 g and 2440 g) and no reported hearing problems or neurological conditions. Mother's average age was 33.0 years (SE = 4.3). Four were unemployed and seventeen had attended college. The mean gestational period was 38.7 weeks (SE = 1.5). In seventeen dyads, the infant was the first child; in the remaining eight dyads, the infant was the second child. All mothers gave informed written consent before their participation in the study and agreed to the videotaping of the social interaction, respecting their privacy and confidentiality, for posterior use for research purposes. All parents gave informed consent of the procedure and the study was approved by the University of Minho ethics committee.

Procedure and Materials

Participants were videotaped while interacting in a child-friendly room. Mother and infant sat on the floor on a soft carpet and mothers were asked to interact, as they naturally would. A camera was placed to capture a side view of the dyad. The semi-structured interaction was composed of three tasks, with a small pause in between. Tasks consisted of: (1) a free-play interaction with toys, (2) a free-play interaction without toys, and (3) a challenging-toy play interaction. Each task had an approximate duration of 3 minutes at 7 months and 5 minutes at 12 months. The experimental design followed the protocol to measure Ainsworth's mother sensitivity and cooperation, and tasks were conducted in a fixed order (1 – free-play with toys, 2 – free-play without toys, 3 – challenging-toy play) from the least stressful to the potentially more stressful interaction. In the free-play task with toys, mothers were instructed to play freely using toys suitable to the infant's age; these were selected from Bayley-III and/or Griffiths 0–2 (7 months – ball, bell, doll, cubes set, cup, mirror, plush bear, rattle toys, squeeze toys, story book, spoon, rubber ring, stacking ring set; 12 months – bell, blanket, cup, doll, mirror, plush bear, rubber ring, stacking pegboard set, squeeze toy, spoon, story book). In the free-play task without toys, mothers were asked to play as they usually do, when there are no toys available. In the play task with a challenging toy, mothers were asked to help their infants play with a challenging toy, selected to be above the infant's developmental level (7 months – a ball that could be squeezed; 12 months – a shape sorter, i.e. a box with holes in the shape of geometric figures where matching 3D objects can be fitted and dropped inside the box). The toys used were only visible to the infant during each respective task. Before every task, the experimenter entered the room, provided general instructions and placed or removed the objects from the floor according to the task. The period when the investigator was in the room was disregarded in the analysis of mother-infant vocalizations. Before the beginning of the procedure, mothers were informed that they could stop the interaction at any time if they considered the infant was uncomfortable or tired due to excessive fretting or crying. To ensure that the infant was in an alert state and more available to perform the tasks, the laboratory visit was scheduled to fit the infant's eating and sleeping patterns.

Design

The study followed a longitudinal design with two dependent variables (duration of overlap and duration of gap), measured for each interaction partner (mother, infant), at two age points (7 and 12 months), and in three different tasks (1 – free-play with toys, 2 – free-play without toys, 3 – challenging-toy play), in a fixed order. We also examined the frequency of gaps and overlaps. Additionally, we should make clear that although one of the conditions is similar to a face-to-face interaction, our procedure does not mimic the more static setup, usually adopted with younger infants, where mother and infant are seated facing each other. Instead, interactions were only restricted to the space, which resulted in interaction moments were mother and infant actually face each other, but also moments were their orientation may vary, throughout the task (e.g. the mother held the infant in her lap).

Coding of Gaps and Overlaps in Turn-transitions

For the coding of mother and infant turn-transitions, vocalizations were first manually segmented by marking their onset and offset, using the ELAN software (Wittenburg et al., 2006). Assistants were trained to segment vocalizations and 100% of each dyad's coding was reviewed by the first author. The resulting segments were then automatically coded, following the on-off binary logic of the AVTA model, the Automated Vocal Transaction Analyzer (Cassotta et al., 1964). AVTA first computes four dyadic states of dialogue: 0 = both participants are silent, 1 = mother vocalizes while infant/child is silent; 2 = infant/child vocalizes while mother is silent; 3 = both participants vocalize simultaneously. From these coded states, segments were further computed using a turn rule, defined by Jaffe and Feldstein's (1970) dialogical approach. This rule gives turn ownership (who holds the floor) to the participant that vocalizes alone until the other participant vocalizes alone. This results in five vocal states: vocalization (V), pause (P), non-interruptive simultaneous speech (NSS), interruptive simultaneous speech (ISS), and switching pause (SP). Two states are of interest in the present study: (1) interruptive simultaneous speech (ISS) marks an overlapping vocalization followed by a change in who has the floor of the conversation; and (2) switching pauses (SP) mark the silences between the vocalization of the partner who has the floor and the vocalization of the partner who gains the floor.

The segments coded as interruptive simultaneous speech (ISS) correspond to the overlaps and the segments coded as switching pauses (SP) correspond to gaps in turn-transitions. Given that the AVTA model is inherently dialogic and we were interested in the individual production of overlaps and gaps, we also take in consideration the direction of turn-transition – i.e., (1) mother to infant; (2) infant to mother. We attribute gaps and overlaps to the partner that gains the turn, rather than determining the ownership of overlaps and gaps by the partner that loses the floor, as in Jaffe and Feldstein's (1970) turn rule.

Analysis of Turn-transitions

To understand the developmental trajectory of turn-transitions, in terms of frequency, we calculated the proportion of overlaps and gaps by task duration (3min, 5min), for each partner (mother, infant) and condition (free-play with objects, free-play without objects, challenging-object play), at 7 and 12 months. In terms of turn-transition duration, we calculated mean and median durations of gaps and overlaps, for each partner and condition, at both age-points. To assess the significance of the changes in duration, we used generalized linear mixed effects modeling and modelled gaps and overlaps separately. The unit of analysis was each individual turn-transition, i.e. each dyad contributed with multiple gaps and overlaps. The distribution of gap and overlap duration was highly skewed (to be expected since this corresponds to cutting in two parts the unimodal distribution of floor-transfer offset, centered close to zero) and no standard transformation made them approximately Normally distributed. We fitted generalized linear mixed models with a Gamma distribution for the response variable, and the canonical inverse function for the link function. The variables that were included in as fixed effects were three categorical predictors. Infant age in months (7, 12 months), task (free-play with objects, free-play without objects, challenging-object play), and partner (mother, infant). The random effects included a random intercept per dyad.

Two best-fit models were selected using a model comparison approach, one with gap duration as the dependent variable, and another with overlap duration as the dependent variable. The final model was selected by starting with a null model that included only the random effect and incrementally adding fixed effects and interaction terms. The effects that were significant according to a likelihood ratio were kept. Models were fitted using the *gmler* function of the *lme4* package (Bates et al., 2012) in R (R Development Core Team, 2012). Statistical inference was based on computing the estimated marginal means and corresponding 95% confidence interval. This was done using the R package emmeans (Lenth, 2021).

Results

Descriptive Statistics

Frequency

After applying the AVTA model to the segmented vocalizations, a total of 5014 turn-transitions were detected, with 2519 turn-transitions attributed to infants, see Table 1. At 7 months there were 745 turn-transitions, from which 76% were gaps and 24% were overlaps. At 12 months a total of 1774 turn-transitions were attributed to infants, 25% were overlaps. There were more turn-transitions initiated by infants in the free-play without toys task (1199) than in the free-play with toys task (610) or the challenging object play task (710). Similarly, mothers produced a total of 2495 turn-transitions. From the 735 turn-transitions assigned to mothers, at 7 months, 76% were gaps. At 12 months, mothers produced 1760 turn-transitions, 27% were overlaps. Mother's turn-transitions were also more for the second task – free-play without toys (1191) – when compared with the first – free-play with toys (601) – and third – challenging object play (703). Overall, the number of turn-transitions and the proportions of gaps and overlaps, followed a similar pattern for mothers and infants.

Table 1

Frequency of Turn-transitions in the Dataset by Partner, Age, and Task

Partner	Age	Free-play with Objects	Free-play without Objects	Challenging Object Play	Total
Infant	7 months	157	354	234	745
	Gap	78%	71%	82%	76%
	Overlap	22%	29%	18%	24%
	12 months	453	845	476	1774
	Gaps	83%	72%	74%	75%
	Overlaps	17%	28%	26%	25%
	Total	610	1199	710	2519
Mother	7 months	154	349	232	735
	Gap	79%	71%	80%	76%
	Overlap	21%	29%	20%	24%
	12 months	447	842	471	1760
	Gap	78%	71%	72%	73%
	Overlap	22%	29%	28%	27%
	Total	601	1191	703	2495
Grand total		1211	2390	1413	5014

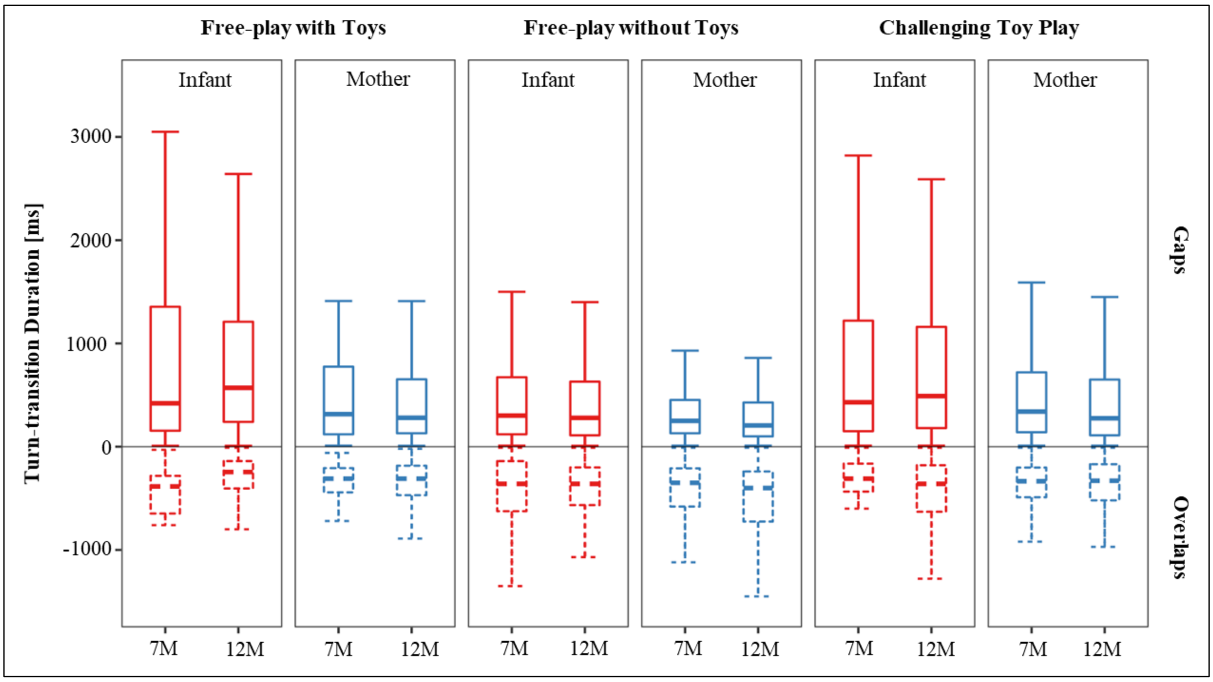
Note. Turn-transition counts are presented for each combination of age, task, and person. Inside each cell of age, task, and person we show the proportion of gaps and overlaps relative to the cell's count; e.g. for infants at 7 months in the free-play with objects task there 157 turn-transitions, and 78% of them were gaps.

Duration

Gap durations were extracted from the switching pause (SP) state in the AVTA model, and assume positive values. Overlap durations were extracted from the interruptive simultaneous speech (ISS) state, but are visualized as negative values for convenience, based on the floor-transfer offset convention. Average and median durations were calculated per task, age point, and direction of turn-transition. At 7 months, gap durations averaged around 1012 ms (Mdn = 420 ms) for infants, and 1000 ms (Mdn = 315 ms), for mothers, in the free-play with toys task; 630 ms (Mdn = 301 ms) for infants, and 508 ms (Mdn = 250 ms) for mothers, in the free-play without toys task; and 1016 ms (Mdn = 430 ms) for infants, and 922 ms (Mdn = 340), in the challenging toy play task. At 12 months, infant's gap mean duration for the free-play with toys increased slightly to around 1106 ms (Mdn = 570 ms), while mother's decreased to 697 ms (Mdn = 280 ms); for the free-play without toys, both infant and mother's gap durations became shorter, averaging 500 ms (Mdn = 279) and 377 ms (Mdn = 206) respectively; and, for the challenging toy play, mother's mean gap duration shortened to 767 ms (Mdn = 275), while infant's gap decreased slightly in average duration to 995 ms, median duration actually increases (Mdn = 490 ms). Overlap average durations, in infants, ranged from around -582 ms to -313 ms, which represents the greatest variation in both infant and adult overlap average durations, corresponding to the change between 7 and 12 months, in free-play with toys task. Figure 3 shows the distribution of gap and overlap duration in the dataset.

Figure 3

Boxplot of the Distribution of Turn-transition Duration for Each Combination of Age, Partner, and Task



Note. Overlap durations are presented with negative values for convenience based on floor-transfer offset convention.

Overlap durations appear to remain stable across age points and direction of turn-transition, with greater variability in the free-play without toys task. Gap duration suggests a general difference between the free-play without toys task, compared with the other two object-oriented conditions, with shorter durations in the free-play without toys task, than in the other tasks. Gap durations also appears to vary differently from one age point to the next, depending on task and direction of turn-transition. While mothers produce shorter gaps, at age point 12 months, in all conditions, infant's gap duration only mirrors this pattern in the free-play without toys task, varying only slightly between age points in the object-oriented conditions.

Correlation of Mean Duration Within-dyad, and Across Age

As a global, descriptive measure of stability across age point and similarity within-dyad, we calculated mean turn-transition duration, collapsing across task. That is, we calculated mean gap duration and mean overlap duration, including all turn-transitions, for age (7 months/12 months) x direction (mother to infant/infant to mother). We further calculated the Pearson correlation between the mean duration values within a dyad. At 7 months and 12 months, the correlation between mother and infant; for the infant, the correlation between duration and 7 months and the duration at 12 months; and, likewise, for the mother.

For mean gap duration, mother-infant dyads were moderately correlated at 7 months, $r(23) = .30$, $p = 0.14$, and strongly correlated at 12 months, $r(23) = .62$, $p < 0.001$. There was some stability between the mother's mean gap duration at the two age points, $r(23) = .30$, $p = 0.14$, but no correlation between the infant's mean gap duration at the two age points, $r(23) = .12$, $p = 0.56$. For mean overlap duration, there was only a strong correlation within the dyad at 12 months, $r(23) = .73$, $p < .001$; all other correlations had absolute values below 0.15 and *p*-values above 0.5.

Modelling of Turn-transition Duration

Modelling of turn-transition duration was done separately for gaps and overlaps.

Gaps

First, we built a base model with gap duration as a dependent variable and a random intercept per dyad, as a random effect. We compared the base model to models including age, task, direction, and their interactions, as fixed effects. The model was significantly improved by adding age, $\chi2(1) = 20.23$, $p < 0.01$; task, $\chi^2(3) = 332.89$, $p < 0.01$; direction, $\chi^2(1) = 38.41$, $p < 0.01$; interaction between age and task, $\chi^2(2) = 12.21$, $p < 0.01$; interaction between age and direction, $\chi^2(1) = 7.85$, $p < 0.01$; and interaction of task and direction, $\chi^2(2) = 8.32$, $p < 0.02$. Table 2 shows the fixed effects for the best model (*log-likelihood* = -2036.65, $N = 3741$).

Table 2

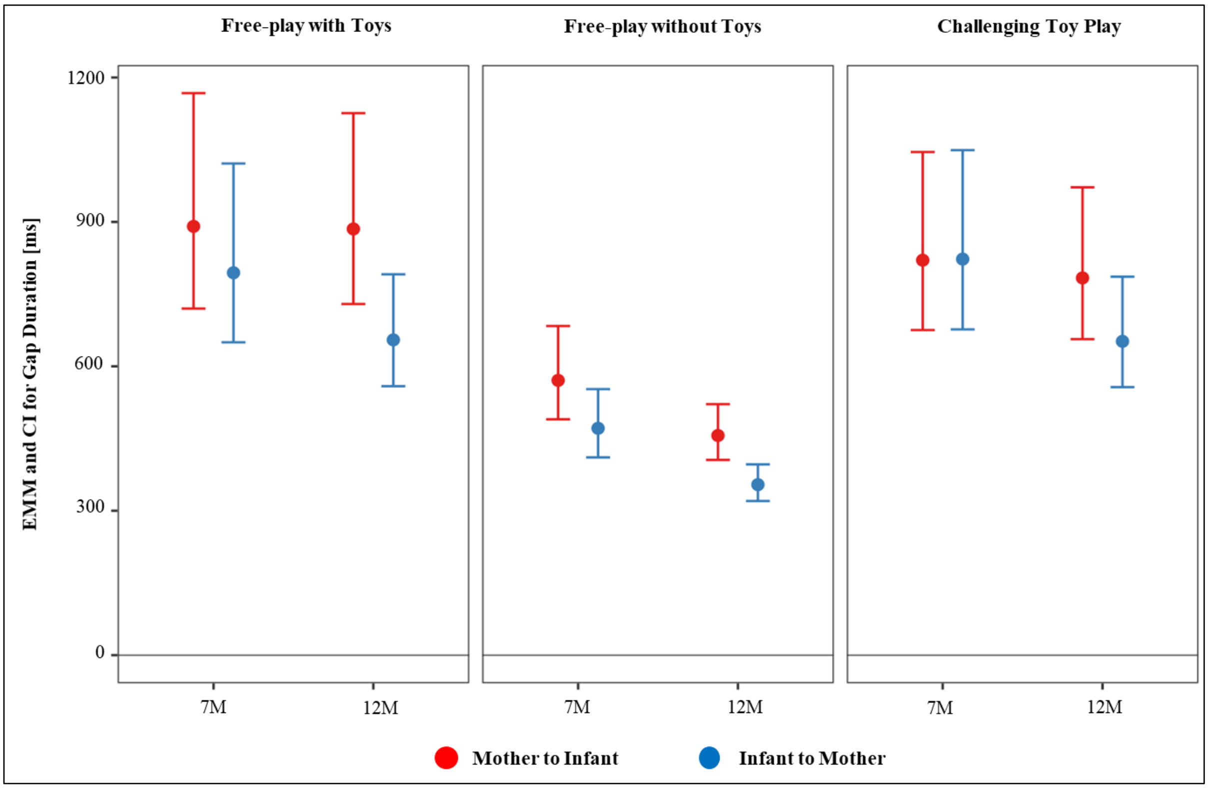
Parameter Estimates, Confidence Interval, and Significance Value for the Fixed Effects in the Model for Gap Duration

Predictors	Estimates	CI	*P*
(Intercept)	4.94	3.93 – 6.21	<0.001
Age (7 months)	0.86	0.82 – 0.91	<0.001
Task (free-play with objects)	0.71	0.67 – 0.76	<0.001
Task (free-play without objects)	1.87	1.72 – 2.03	<0.001
Direction (mother to Infant)	0.86	0.82 – 0.91	<0.001
Age (7 months) x Task (free-play with objects)	1.08	1.02 – 1.16	0.014
Age (7 months) x Task (free-play without objects)	0.87	0.81 – 0.94	0.001
Age (7 months) x Direction (mother to infant)	1.07	1.02 – 1.12	0.005
Task (free-play with objects) x Direction (mother to infant)	1.02	0.95 – 1.08	0.628
Episode (free-play without objects) x Direction (mother to	0.9	0.84 – 0.98	0.012

Figure 4 plots the estimated marginal means with confidence intervals for gap durations, showing a clear difference between the free-play without toys interaction condition and the object-oriented interaction conditions, which can be confirmed by post-hoc pairwise comparisons. Mean duration for the free-play without toys is smaller than the free-play with toys task (z = -14.05, $p < 0.001$) and the challenging toy play task (z = 13.74, $p < 0.001$). The interaction plot (Figure 4) also reveals that the differences between conditions remain at both age points. Moreover, it suggests that the age effect is stronger for the face-to-face interaction condition, with pairwise comparisons confirming that, only for the free-play without toys, the within-task contrast between age points is significant (z = -5.26, $p <$ 0.01). Although, the direction effect suggests a difference between infant and mother gap durations, this difference is not evident in all conditions; only for the first (z = -3.61, $p < 0.01$) and second task (z = -4.56, $p < 0.01$) the within-task pairwise comparisons between direction of turn-transition show significant contrasts.

Figure 4

Predicted Mean Duration and Respective 95% Confidence Interval for Gaps in Turn-transitions



Overlaps

For examining potential differences in overlaps, we built another base model using overlap duration as dependent variable, and a random intercept per dyad as random effect, which we compared to models including age, task, and direction. The model was significantly improved by adding task, $\chi^2(2) = 22.81$, $p < 0.01$; but not by adding age, $\chi^2(1) = 0.01$, $p = 0.93$, nor direction, $\chi^2(1) = 1.58$, $p = 0.21$. Table 3 shows the fixed effects for the final model (*log-likelihood* = -215.49, *N* = 1273).

Table 3

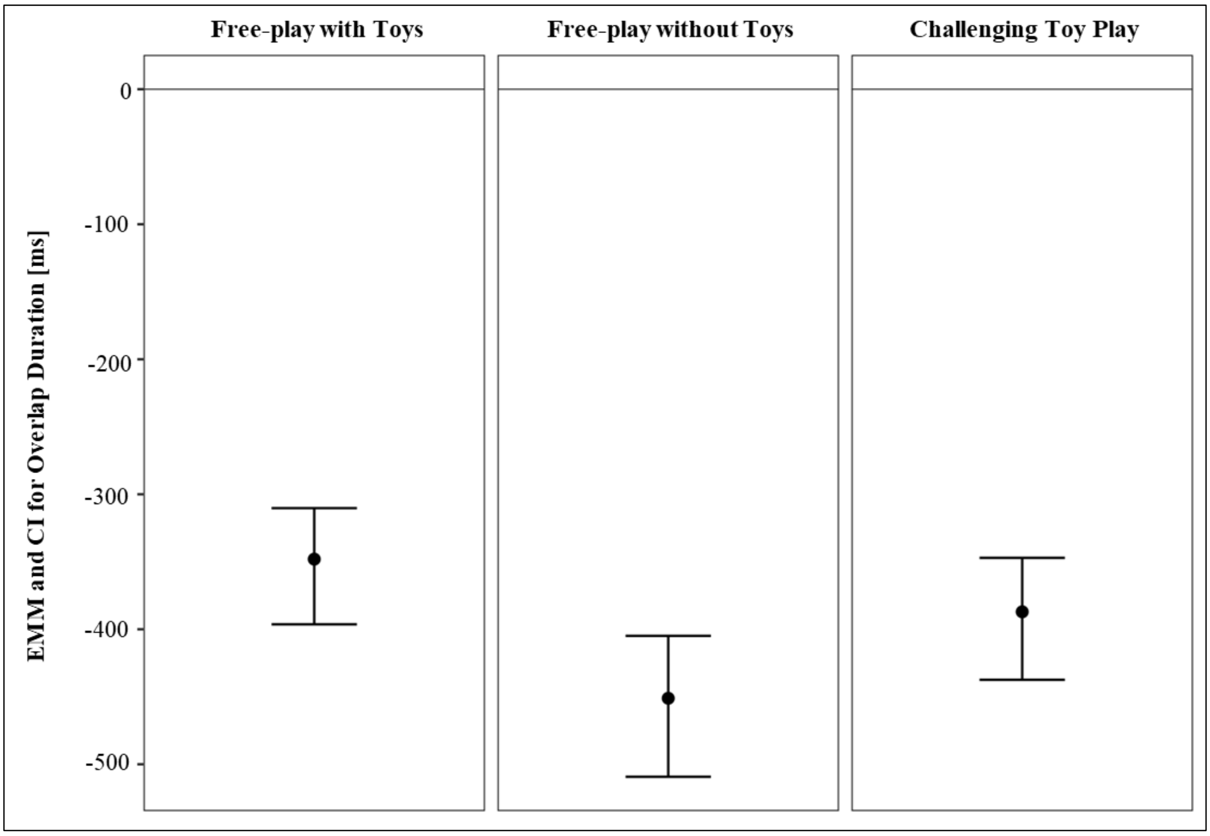
Parameter Estimates, Confidence Interval, and Significance Value for the Fixed effects in the Model for Overlap Duration

Predictors	Estimates	CI	*P*
Intercept	12.91	10.09 – 16.52	<0.001
Task (free-play with objects)	1.37	1.12 – 1.68	0.003
Taks (free-play without objects)	0.71	0.62 – 0.82	<0.001

Since only the addition of task significantly improved the model, there is only an overall effect of task on overlap duration, which does not distinguish between direction of turn-transition or age point. Figure 5 plots the estimated marginal means and confidence intervals for the best model. Post-hoc pairwise comparisons confirm that the free-play without toys is significantly different from free-play with objects ($z = 4.13$, $p < 0.01$) and challenging object play ($z = -3.00$, $p < 0.01$), but that the object-oriented tasks do not differ significantly ($z = 1.62$, $p = 0.24$). Unlike task differences in gap duration, the difference between tasks in overlap duration reveals that in fact overlaps are significantly longer in the face-to-face interaction task.

Figure 5

Predicted Mean Duration and Respective 95% Confidence Interval for Overlaps in Turn-transitions



Note. Although modelled with positive values, overlap durations are presented with negative values for convenience, based on the floor-transfer offset convention.

Discussion

It is within the ecological niche of early interactions that language and the ability to converse develops into what is recognizably the most prevalent organization in human conversation, the turn-taking system (Levinson & Torreira, 2015).

Hypotheses Versus Results

Overall, the results support our expectation that turn-transitions are different whether object manipulation is present or absent in the interaction. Differences between the face-to-face and the object-oriented conditions are evident when comparing the prevalence of turn-transitions – the frequency in the free-play without toys task is almost the double of the other two tasks. Although median durations were generally shorter compared to the findings in Hilbrink et al. (2015), we have also found evidence that supports our hypotheses regarding turn-transition duration and developmental trajectory.

For gap duration, we have shown that, indeed, object-oriented conditions produce significantly longer gaps than the face-to-face condition (Hypothesis 1), which was true for infants and mothers, at both age points. As expected, only for the object-oriented conditions, infant gap durations approaches the values reported in Hilbrink et al. (2015) for 9 months (~1100 ms) and 12 months (~975 ms). Infant gap durations, in the face-to-face condition, are higher than the average floor-transfer offset of adult conversations (~200 ms), but the floor-transfer offset includes overlaps in the calculation and, more importantly, infants' gap durations are substantially closer to the adult-adult average floor-transfer offset, than to their own gap durations, in the object-oriented conditions.

Regarding the developmental trajectory of gap durations, results support our conjecture (Hypothesis 2) that, for object-oriented conditions, infant gap durations would be similarly affected by the presence of object manipulation. We found no significant differences in average duration, between 7 and 12 months. For the face-to-face condition, our results suggest that, as conjectured (Hypothesis 3), there is a developmental decrease in infant gap duration, when objects are removed from the interaction. Interestingly, this decrease appears to follow the descending trajectory of mothers' average gap durations that were shorter, at 12 months, for all tasks.

Finally, concerning infant overlap durations, although there was a general dyadic difference between object-oriented and face-to-face conditions, no age or direction effect was found. This supports our expectation (Hypothesis 4) that there are no significant developmental differences in infant overlap duration, with average durations only slightly below the median durations reported in Hilbrink et al.

(2015) for 9 and 12 months. Interestingly, it also suggests that the presence or absence of object play can have, nonetheless, an impact on overlap duration.

General Discussion

The evidence presented in this study has both methodological and theoretical implications.

On the methodological side, turn-taking research with children (Casillas et al., 2016) and adults (Levinson & Torreira, 2015) examines verbal exchanges during conversation and in some studies, considers only close-ended questions and answers (Berninger & Garvey, 1981). This is not the case with non-verbal infants, whose proto-conversations are mainly studied from the non-verbal vocal exchanges that occur in the context of playful mother-infant interactions (Bateson, 1975; Beebe et al., 1988; Hilbrink et al., 2015). To study the developmental trajectory of turn-taking with infants in playful interactions, researchers should take into consideration the developmental progression of play (Williams, 2003), as well as the infant's ability to engage in certain types of play. Studies must adapt the setup and materials to the infant's developmental level. This kind of adaptations fall most commonly under the category of face-to-face interaction – with younger infants that cannot support their body are seated face-to-face to their mothers; or object-oriented interactions – with older infants, that can support their own body weight, move around the setup, manipulate objects and engage in joint-attention.

The differences highlighted by our research alerts us that, although it is essential to adapt the setup and materials to infant's development, it is also crucial to understand if those adaptations are indeed equivalent in regards to the object of study – turn-taking. What we have found is that assessing turn-transition durations in the context of object play is not the same as examining the same parameters when removing objects from the interaction. Longitudinal studies that encompass several age points and the developmental progression of play, should attend to the experiment design by possibly controlling for the presence or absence of toys in early face-to-face interactions, or by introducing also free-play tasks without toys, and compare turn-transitions throughout time when toys are present or absent.

Theoretically, the evidence we present here also re-opens the discussion on the developmental trajectory of turn-taking, specifically when examining face-to-face interactions.

As we have seen, in the most comprehensive longitudinal study of infant turn-transition duration to date (Hilbrink et al., 2015), gaps and overlaps vary differently over time. While gap duration doubles between 5 (~500 ms) and 9 (~1100 ms) months, it follows a slowly descending trajectory from 12 (~975 ms) to 18 (~700 ms) months. Overlaps vary the most from 3 to 5 months, and appear to remain stable

(~575 ms) from 9 months onwards. This has been interpreted as evidence that gap and overlap duration follow different developmental trajectories, and that gap duration, starting from 9 months, appear to slow down due to the influence of language acquisition. Results in our study suggest that, indeed, gap and overlap duration may follow different developmental trajectories, but hint to alternative explanations for longer gap durations.

In our study, infant's median turn-transition durations were overall shorter than those reported in the previous research, for 9 and 12 months (Hilbrink et al., 2015). Average gap durations, though, were closer to the median gap durations in that study, for the object-oriented tasks than for the face-to-face interaction. We have been interpreting this proximity in gap duration only for the object-oriented conditions so far, by arguing that these tasks may be closer to the conditions of that study, where object manipulation was allowed. Given that infant gap durations are significantly shorter in the face-to-face interaction, we suggest that we have found evidence that the presence of object manipulation has an effect in the duration of gaps between 7 and 12 months; and that the development of infant's ability for active object exploration (Ruff, 1984) and to engage in joint-attention and joint-action (Bakeman & Adamson, 1984; Carpenter et al., 1998), may explain in part the differences between younger and older infants, in Hilbrink et al. (2015).

Although we have not examined gap durations at other age points, by introducing an age point – 7 months – and having a comparable age point we can suggest at least some extrapolations. We did not find a significant difference between 7 and 12 months in the object-oriented conditions, similarly, in Hilbrink et al. (2015), there was no significant difference between 9 and 12 months, but actually a significant increase in gap duration from 5 to 9 months, followed by a significant decrease between 12 and 18 months. We suggest, then, that gap duration is already, at 7 months, similarly long to the gaps produced at 9 and 12 months. Interpreting such an increase in gap duration, as a function of the emergence of language, seems unlikely at 7 months.

Looking at the face-to-face interaction though, we see a significant decrease in gap duration between 7 and 12 months, which although significantly different in direction, is quite similar to the pattern in time presented by mothers, in every condition. The shorter durations and the significant decrease, suggest that, when objects are removed from the interaction, we can observe that infants can already produce, at 7 months, gaps that are much closer to those expected in adult conversations, and that they can even reduce their duration, by 12 months. Together with the similarities in the tendency between infant and mothers, we may suggest that the mother-infant dyads are becoming more efficient at predicting turn-transitions. This is consistent with the interaction engine hypothesis (Levinson, 2006) that suggests that the foundations for turn-taking, such as turn-timing sensitivity, are already present in pre-verbal infants, enabling them to have tighter turn-transitions in the second half of their first year of life.

Finally, our study has some limitations. The order of conditions was fixed, with task (1) free-play with toys, always followed by task (2) free-play without toys, and finally by task (3) challenging object play. Although we were not able to counterbalance the tasks, we were able to show that the differences between conditions are captured not only when changing from (1) an object-oriented task to (2) a face-to-face task, but also when changing from (2) a face-to-face task to (3) an object-oriented task. Another limitation of our study is that it only examines two age points, and only one is directly comparable with previous research. Given the age points in our study, we cannot discard that an increase in gap duration may actually occur at 9 months, although we also have no evidence that, at 12 months, language emergence is producing longer gap durations.

Conclusion

We have shown that, by 7 months, there is already a significant difference in infant's gap durations, between face-to-face and object-oriented interactions, and that significant differences between age points, that may indicate a direction in the developmental trajectory, are only detected in the face-to-face interaction condition. Through this, we also offer a complementary explanation for longer gap durations in previous studies, which appear to be more prevalent in object-oriented interactions than in face-to-face interactions, between 7 and 12 months. Furthermore, it alerts developmental researchers that differences in the context of data collection should be accounted when designing experiments and analysing data from turn-taking developmental research. It also reframes the discussion on the foundational structure of turn-taking (Levinson, 2006), that predates the process of language acquisition.

References

Adolph, K. E., & Hoch, J. E. (2019). Motor development: embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, *70*(September 2018), 141–164. https://doi.org/10.1146/annurev-psych-010418-102836

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*(4), 1278. https://doi.org/10.2307/1129997

Bakeman, R., Adamson, L. B., Konner, M., & Barr, R. G. (1990). !Kung infancy: The social context of object exploration. *Child Development*, *61*(3), 794. https://doi.org/10.2307/1130964

Bates, D., Maechler, M., & Dai, B. (2012). *Lme4: Linear mixed effects models using s4 classes.*

Bateson, M. C. (1971). The interpersonal context of infant vocalization. *Quarterly Progress Report of the Research Laboratory of Electronics*, *100*, 170–176.

Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, *263*(1), 101–113. https://doi.org/10.1111/j.1749-6632.1975.tb41575.x

Beebe, B., Alson, D., Jaffe, J., Feldstein, S., & Crown, C. (1988). Vocal congruence in mother-infant play. *Journal of Psycholinguistic Research*, *17*(3), 245–259. https://doi.org/10.1007/BF01686358

Berninger, G., & Garvey, C. (1981). Questions and the allocation, construction, and timing of turns in child discourse. *Journal of Psycholinguistic Research*, *10*(4), 375–402. https://doi.org/10.1007/BF01067165

Bertenthal, B. I., & Boyer, T. (2015). The development of social attention in human infants. In A. Puce & B. I. Bertenthal (Eds.), *The many faces of social attention: Behavioral and neural measures* (pp. 21–65). Springer International Publishing. https://doi.org/10.1007/978-3-319-21368-2

Bigelow, A. E., & DeCoste, C. (2003). Sensitivity to social contingency from mothers and strangers in 2-, 4-, and 6-month-old infants. *Infancy*, *4*(1), 111–140. https://doi.org/10.1207/S15327078IN0401_6

Bloom, K. (1977). Patterning of infant vocal behavior. *Journal of Experimental Child Psychology*, *23*(3), 367–377. https://doi.org/10.1016/0022-0965(77)90032-7

Brazelton, T. B., Koslwski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). Willey-Interscience.

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language*, *2*(1), 1–19.

https://doi.org/10.1017/S0305000900000866

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i. https://doi.org/10.2307/1166214

Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn-taking, timing, and planning in early language acquisition. *Journal of Child Language*, *43*(6), 1310–1337. https://doi.org/10.1017/S0305000915000689

Casillas, M., & Frank, M. (2013). The development of predictive processes in children's discourse understanding. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 299–304. http://langcog.stanford.edu/papers/CF-cogsci2013.pdf

Cassotta, L., Feldstein, S., & Jaffe, J. (1964). AVTA: A device for automatic vocal transaction analysis. *Journal of the Experimental Analysis of Behavior*, *7*(1), 99–104. https://doi.org/10.1901/jeab.1964.7-99

Condon, W. S., & Sander, L. W. (1974). Synchrony demonstrated between movements of the neonate and adult speech. *Child Development*, *45*(2), 456. https://doi.org/10.2307/1127968

Elias, G., & Broerse, J. (1996). Developmental changes in the incidence and likelihood of simultaneous talk during the first two years: A question of function. *Journal of Child Language*, *23*(1), 201–217. https://doi.org/10.1017/s0305000900010151

Elias, G., Hayes, A., & Broerse, J. (1986). Maternal control of co-vocalization and inter-speaker silences in mother-ineant vocal engagements. *Journal of Child Psychology and Psychiatry*, *27*(3), 409–415. https://doi.org/10.1111/j.1469-7610.1986.tb01842.x

Fogel, A. (1977). Temporal organization in mother-infant face-to-face interaction. *Studies in Mother-Infant Interaction*, *January 1977*, 119–152.

Garvey, C., & Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse Processes*, *4*(1), 27–57. https://doi.org/10.1080/01638538109544505

Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, *15*, 221–235. https://doi.org/10.1017/S0305000900012344

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology*, *6*(September), 1–12. https://doi.org/10.3389/fpsyg.2015.01492

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), i–viii, 1–132. http://www.ncbi.nlm.nih.gov/pubmed/11428150

Jaffe, J., & Feldstein, S. (1970). *Rhythms of dialogue*. Academic Press.

Jasnow, M. D., & Feldstein, S. (1986). Adult-like temporal characteristics of mother-infant vocal interactions. *Child Development*, *57*(3), 754–761. http://www.jstor.org/stable/1130352

Kaye, K., & Fogel, A. (1980). The temporal structure of face-to-face communication between mothers and infants. *Developmental Psychology*, *16*(5), 454–464. https://doi.org/10.1037/0012-1649.16.5.454

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means. R package version 1.5.5-1.* https://cran.r-project.org/package=emmeans

Levinson, S. C. (2006). On the human interactional engine. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality. Culture, Cognition and Interaction* (pp. 39–69). Berg.

Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 103–130). Willey-Blackwell.

Levinson, S. C. (2019). Interactional foundations of language: The interaction engine hypothesis. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 189–200). MIT Press.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*(JUN), 1–17. https://doi.org/10.3389/fpsyg.2015.00731

Lourenço, V., Coutinho, J., & Pereira, A. F. (2021). Advances in microanalysis: Magnifying the social microscope on mother-infant interactions. *Infant Behavior and Development*, *64*(April), 101571. https://doi.org/10.1016/j.infbeh.2021.101571

Masataka, N. (1993). Effects of contingent and noncontingent maternal stimulation on the vocal behaviour of three- to four-month-old Japanese infants. *Journal of Child Language*, *20*(2), 303–312. https://doi.org/10.1017/S0305000900008291

Miguel, H. O., Gonçalves, Ó. F., & Sampaio, A. (2020). Behavioral response to tactile stimuli relates to brain response to affective touch in 12-month-old infants. *Developmental Psychobiology*, *62*(1), 107–115. https://doi.org/10.1002/dev.21891

Miguel, H. O., Lisboa, I. C., Gonçalves, O. F., & Sampaio, A. (2019). Brain mechanisms for processing discriminative and affective touch in 7-month-old infants. *Developmental Cognitive Neuroscience*, *35*(October 2017), 20–27. https://doi.org/10.1016/j.dcn.2017.10.008

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social Perception in Infants* (pp. 177–197). Ablex

Publishers.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science*, *2*(2), 164–173. https://doi.org/10.1111/1467-7687.00065

R Development Core Team. (2012). *R: A language and environment for statistical computing.* r foundation for statistical. http://www.r-project.org/

Rochat, P., Neisser, U., & Marian, V. (1998). Are young infants sensitive to interpersonal contingency? *Infant Behavior and Development*, *21*(2), 355–366. https://doi.org/10.1016/S0163-6383(98)90012-1

Ruff, H. A. (1984). Infants' manipulative exploration of objects: Effects of age and object characteristics. *Developmental Psychology*, *20*(1), 9–20. https://doi.org/10.1037/0012-1649.20.1.9

Rutter, D. R., & Durkin, K. (1987). Turn-taking in mother-infant interaction: An examination of vocalizations and gaze. *Developmental Psychology*, *23*(1), 54–61. https://doi.org/10.1037//0012-1649.23.1.54

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Serra, J., Miguel, H., Moura, A. A., Sampaio, A., & Pereira, A. F. (2020). The effect of play task on maternal touch patterns when interacting with their 12 months-old infants: An exploratory study. *Infant Behavior and Development*, *59*(March), 101438. https://doi.org/10.1016/j.infbeh.2020.101438

Stern, D. N. (2002). *The first relationship: Infant and mother* (2nd ed.). Harvard University Press.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies*, *7*(2), 251–271. https://doi.org/10.1075/is.7.2.08str

Thorgrimsson, G. B. (2014). *Infants' understanding of communication as participants and observers*. [Unpublished doctoral dissertation]. Radbound University Nijmegen.

Tronick, E. D., Als, H., & Brazelton, T. B. (1977). Mutuality in Mother-Infant Interaction. *Journal of Communication*, *27*(2), 74–79. https://doi.org/10.1111/j.1460-2466.1977.tb01829.x

Tronick, E. Z., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Williams, E. (2003). A comparative review of early forms of object-directed play and parent-infant play in typical infants and young children with autism. *Autism*, *7*(4), 361–377. https://doi.org/10.1177/1362361303007004003

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of the 5th International Conference on Language Resources and Evaluation, LREC 2006*, *August*, 1556–1559.

CHAPTER III – STUDY 2[2]

---

[2] A version of this chapter was submitted to the Wiley-Blackwell peer-reviewed journal *Developmental Science*:

Lourenço, V., Serra, J., Coutinho, J., & Pereira, A. F. (2022). *Turn-taking in free-play interactions: A Cross-sectional study from 3 to 5 years*. [Manuscript submitted for publication]. School of Psychology, University of Minho.

Abstract

Turn-transition timing in childhood has been examined by measuring response latency – that aggregates gap and overlap duration – in turn-transitions contingent to specific semantic categories. This contrasts with studies in infancy where the whole spectrum of temporal contingent vocalizations are examined, and gap and overlap duration is analyzed independently. We propose using the latter approach to investigate the continuities between infancy and childhood. In a cross-sectional design, we analyzed the vocalizations of 44 mother-child free-play interactions, ranging from three to five years of age. Frequency and duration were measured for gaps and overlaps, independently, and as an aggregated measure – floor-transfer offset (FTO). The effects of child's age and direction of turn-transition (child, mother) were assessed using generalized linear mixed modelling for each dependent variable (DV: FTO, gaps, overlaps). Although there was a slight increase in FTO and gap duration across ages, no significant effect of age was found for any of the DVs. There was an effect of turn-transition direction, for FTO and gap durations, but not for overlap duration. Children-initiated transitions produced significantly longer FTO and gap durations than their mothers, but had similarly timed overlaps. Results suggest that gaps and overlaps still have different developmental trajectories throughout childhood; and that overlap duration converges to adult standards, at least, by 3-years-old. Methodologically, we demonstrated the relevance of using complementary metrics (FTO, gap, overlap) to understand the developmental trajectories of turn-taking; and that examining all temporally contingent vocalizations can provide a valid and more inclusive measure of turn-transition duration in childhood.

Keywords: turn-taking; mother-child interaction; floor-transfer offset; gaps; overlaps

Turn-Taking in Free-play Interactions: A Cross-Sectional Study from 3 to 5 Years

Turn-taking is the predominant structure that organizes human conversation (Levinson, 2006). It is characterized by an alternation pattern, where each partner speaks mostly one at a time and floor transitions occur in a way that avoids prolonged silences and superposition (Sacks et al., 1974). This temporal coordination, described as the minimal-gap minimal-overlap phenomenon (Levinson & Torreira, 2015), acts as a glue that holds turns together (Casillas, 2014).

Background

Conventionally, turn-transitions are measured by the floor-transfer offset (FTO), a metric that groups gaps and overlaps in the same time scale. In adult-adult conversations, gap durations averages around 200 ms, and FTO present an unimodal and slightly asymmetrical distribution, detectable across different cultures (Levinson & Torreira, 2015; Stivers et al., 2009). This level of coordination implies that for turn-transitions to occur at such speeds, one must be able to predict where another's turn will end, and simultaneously prepare what to say when the floor switches (Levinson, 2013; Sacks et al., 1974).

Levinson's (2006) interaction engine hypothesis suggests that the predictive abilities of turn-taking, such as a sensitivity to turn timing, and the ability to predict the communicative intentions of others, are aspects of a foundational interactional system, where turn-taking is already prevalent, and on which languages builds up. We will refer to the processes that subserve social coordination, in particular turn-taking, as the interaction system, following Levinson's terminology.

Although turn-taking and language appear to be greatly interconnected, there is evidence that the two have a distinct developmental trajectory, and that the process of integration between the foundational interaction system and language may impose a slowdown in turn-transition timing before converging to the minimal-gap minimal-overlap adult standard (Hilbrink et al., 2015; Levinson, 2019).

Turn-Taking Development

Observational studies using microanalysis (Lourenço, Coutinho, et al., 2021) have shown, that infants are able to engage in contingent interaction across different modalities of communication (Brazelton et al., 1974; Fogel, 1977; Kaye & Fogel, 1980; Stern, 1971; Tronick et al., 1977). As early as 1.5 months, infant's engage in proto-conversations with their mothers, and these vocal exchanges exhibit the alternation properties of turn-taking, long before language acquisition (Bateson, 1975; Jasnow & Feldstein, 1986).

Experimental studies, manipulating the temporal contingency of caregiver-infant interactions to create unresponsive interactions – an adaptation of the still-face paradigm (Tronick et al., 1978) –, have shown indeed that infants, by at least 3 months of age, are sensitive to alterations in contingency (Murray & Trevarthen, 1985; Striano et al., 2006). Other experimental studies tracking ocular movements have shown that infants, at least from 12 months, are also able to accurately predict turn-transitions before they occur (Casillas & Frank, 2013) and distinguish between non-speech and speech vocalizations (Thorgrimsson, 2014).

Earlier work on the analysis of the temporal properties of turn-transitions throughout development has produced inconsistent results (Bateson, 1975; Beebe et al., 1988; Elias et al., 1986; Garvey & Berninger, 1981; Jasnow & Feldstein, 1986). Few longitudinal studies, small samples, the use of different metrics, and the focus on only gaps or overlaps, are some of the limitations that contributed to an unclear picture of the developmental trajectory of turn-taking. More recently, longitudinal and cross-sectional studies have begun to provide finer-grained analysis of gaps and overlaps, with larger samples, across several age points (Casillas et al., 2016; Hilbrink et al., 2015; Lourenço, Pereira, et al., 2021; Stivers et al., 2018).

Hilbrink et al. (2015) studied the durations of gaps and overlaps in free-play interactions with 12 mother-infant dyads, at 3, 4, 5, 9, 12, and 18 months. Infants' median gap durations ranged around 600 ms from 3 to 5 months, increased at 9 months to around 1,100 ms and decreased over time to around 700 ms at 18 months. Mothers' gap duration followed a similar trajectory but were shorter in duration. Infants' median overlap durations has very little change, slightly varying around 500-600 ms from 3 to 5 months, and than stabilizing at 575 ms, from 9 to 18 months. Significant age and age by person (mother or infant) interaction effects were found for both overlap and gap duration.

Results were interpreted as evidence that gaps and overlaps have a different developmental trajectory: while overlaps have a fairly stable trajectory, closer to adult standards; gaps have a substantial increase at 9 months, and then progressively decrease (Hilbrink et al., 2015). Moreover, the longer gap durations at 9 months were interpreted as evidence of a slowdown effect in turn-taking timing due to the integration of the interaction and language systems; and the gradual decrease in gap duration over time, as evidence of a tendency towards the minimal gap standard of adult conversations (Hilbrink et al., 2015).

Lourenço, Pereira, et al. (2021) also studied the durations of overlaps and gaps, but at 7 and 12 months, in 25 mother-infant dyads. The effect of face-to-face and object-oriented interactions in turn-transition duration was also considered. Infants' median gap durations only slightly increased over time for the object oriented tasks, from 420 ms *(M* = 1,012 ms) to 570 ms *(M* = 1,106 ms) in the free object-play task, and 430 ms to 490 ms in the challenging-toy task. Inversely, median gap durations slightly decreased on the face-to-face interaction, from 301ms *(M* = 630 ms) to 279 ms *(M* = 500 ms). Mothers' gap duration consistently followed a decreasing trajectory between age points for all tasks, with shorter times than infants. Significant effects of age, task, direction of turn-transition, age and task interaction, and age and direction interaction, were found for gap duration. Infant gaps where significantly longer than their mothers. When toys were removed from the interaction, infants not only could produce shorter gaps than in object-oriented tasks, but also followed the same descending trajectory as their mothers, with shorter gaps at 12 months compared to 7 months. In contrast, infants' overlap durations did not show a consistent change between tasks, direction or age point; average duration ranging from 313 ms to 582 ms. Only a general effect (independent of partner or age) of task was found for overlaps.

Authors interpreted these results as evidence that infant turn-transition durations, in the second half of the first year, may be considerably shorter than those reported in Hilbrink et al. (2015). Of importance, longer gap durations in this period may reflect the predominance of object-play, which infants gradually engage more often in that period, and the possible interference of the ongoing motor developments in the flow of the interaction (Lourenço, Pereira, et al., 2021). Moreover, if gap duration is smaller when objects are removed from the interaction , that may be an indication that the expected slowdown due to integration between interaction and language systems may be further down the developmental line, when linguistic processing becomes more complex (Lourenço, Pereira, et al., 2021).

Casillas et al. (2016) analyzed the response latency in pair-adjacent (question-responses pairs) turn transitions selected from the naturalistic conversations of 5 children with their caregivers, at six age points from 1;8 (20 months) to 3;5 (41 months) years old. Six combinations of question type (yes/no, wh-) and levels of answer complexity (3 levels) were analyzed. Similar to the FTO convention, response latencies aggregated both overlap and gap durations in the same scale – in which the former assume negative latency values, and the latter positive values. When considering all question-answer combinations, no age effect was found, with a median response latency of 625 ms, across age points. Meanwhile, there was an effect for the level of answer complexity, with longer latencies for more complex responses. When analyzing only yes/no question-answer pairs though, Casillas et al. (2016) found a significant effect of age, response complexity and the interaction between both predictors. Children's response latencies significantly decreased from the first age point (*Mdn* = 651ms) to the last age point (*Mdn* = 469 ms), and simpler responses had a significantly shorter latency than, more complex responses.

The authors suggest that the development of children's language abilities may actually obscure the detection of developmental patterns in turn-taking timing (Casillas et al., 2016). Which is coherent with the idea that although turn-transition durations should be converging to the adult standard (Levinson & Torreira, 2015) the actual development is prolonged over time and non-linear, accommodating the progressive development of children's linguistic abilities (Ervin-Tripp, 1979; Garvey & Berninger, 1981; Hilbrink et al., 2015; Lindsay et al., 2019).

Further evidence of a slow progression can be found in older children. Building on previous research with adults (Stivers et al., 2009), Stivers et al. (2018) studied the response latency but with 4 to 8-years-olds. Using a cross-sectional design, this study examined a large spectrum of pair-adjacent turn-transitions from the naturalistic triadic interactions of 95 school-aged, considering morpho-syntactic and pragmatic aspects. In the final analysis, children were grouped in the 4-5 years range and in the 6-8 years range. Younger children had a modal offset of 400 ms, a median of 500 ms, a mean of 636 ms, and a standard deviation of 687 ms. Older children had a modal offset of 300 ms, a median of 400 ms, a mean of 515 ms, and a standard deviation of 654 ms. Turn-transition durations in both groups were longer than in adults, but indicative of a slight age improvement in turn-taking timing. Additionally, results revealed distributional differences between the type of pair-adjacent transitions consistent with adult patterns. Answer responses were significantly faster than non-answer responses; interjections were significantly faster than other answer types; and confirmations were significantly faster than disconfirmations.

Additionally, in a recent meta-analysis (Nguyen et al., 2022) considering only the gap durations of 26 turn-taking studies with typical and atypical populations, with ages ranging from 0 to 96 months, results predicted an ascending trajectory in gap duration, at least until 40 months, when the trajectory may start exhibit a gradual descending tendency.

Research Objectives

Within the psycholinguistic study of turn-taking there is a striking methodological difference between pre-verbal and verbal infants. While in the study of pre-verbal infants the whole interaction, regardless of the content of vocal exchanges, is measured and analyzed (Hilbrink et al., 2015), in studies of verbal children only pair-adjacent (question-response) transitions are analyzed (Casillas et al., 2016; Garvey & Berninger, 1981). This reflects not only the evident development of children's linguistic abilities, but also a theoretically-driven shift in the focus of turn-taking research, from temporal contingency exclusively, to temporal contingency within semantically contingent transitions (Casillas, 2014). From a microanalytic perspective though, temporal contingency in the whole interaction is relevant and provides a reliable measurement of interpersonal coordination (Jaffe et al., 2001).

Another relevant methodological difference in the study of turn-taking between preverbal and verbal children is how turn-transitions are measured. In pre-verbal infants gaps and overlaps have been measured independently and shown to have different developmental trajectories (Hilbrink et al., 2015; Lourenço, Pereira, et al., 2021). With verbal children, gaps and overlaps are grouped and analyzed in the same time scale – response latency, where overlaps assume a negative value, and gaps a positive (Casillas et al., 2016). Similarly to adult FTO measurements, overlaps are assumed to be a measure of infants' effort to anticipate turn-transition, and as such are measured as if they are negative gaps (Levinson & Torreira, 2015). Although the aggregation makes comparisons with adult turn-taking easier, it also prevents us to understand the developmental trajectory of each type of turn transition.

With theses differences in mind and the limitations that they pose to the comparison between studies and the tracing of the developmental trajectory of turn-taking, we designed a cross-sectional study to understand the development of turn-transition duration between 3;3 and 5;10 years-old. We asked 44 mother-child dyads to engage in a 10 minutes free-play task, where toys were available and conversation was permitted.

We measured all vocal exchanges of the dyads and analyzed the trajectory of turn-transitions across age. First, by aggregating gaps and overlaps in the same measure – floor-transfer offset (FTO) –, where gaps assume positive values and overlaps negative. Then, by analyzing gaps and overlaps independently. The effects of child's age and direction of turn-transition direction (child, mother) in turn-transition duration were tested, and when directional differences were detected, the effect of age was tested independently for the turn-transitions initiated by children or their mothers.

By juxtaposing the results of the FTO analysis to the analyses of gaps and overlaps, we expect to illustrate the benefits of using that approach to pinpoint the contribution of each dimension (gaps, overlaps) to turn-transition timing. Additionally, by using all temporally contingent vocalizations to measure turn-transition duration, we expect to get a more inclusive measure of children's turn-transition timing. One that is not restricted by semantic contingence and may be used as reference for children's turn-timing, both when comparing between infant and children turn-transition durations, and, when comparing to specific levels of linguistic complexity.

Given that the developmental trajectory of turn-transition timing may be non-linear and dependent on linguistic complexity, we do not expect that a more inclusive (all vocalizations) and less granular (no linguistic distinction) approach will detail the potential of possible developmental directions between levels of complexity. We expect, however, to get a better understanding of the developmental trajectory of gaps and overlaps throughout childhood, and their contribution to overall turn-transition timing.

Method

Participants

The data reported here is part of a larger research project – DYNATURNTAKE - studying vocal and motor coordination in parent-child and stranger-child play interactions. In the present study we examined a subset from that project consisting of 44 mother-children dyads. Children's age ranged from 3 years and 3 months (3;3) to 5-years and 10-months (5;10) old. Mothers and their infants were recruited in preschools and daycare centers in Guimarães, Portugal. All children were typically developing infants and no hearing problems or neurological conditions were reported. All mothers gave informed written consent for the procedure, in advance, and agreed to the videotaping of the social interaction, respecting their privacy and confidentiality, for posterior use for research purposes. The study was approved by the University of Minho ethics committee.

Procedure and Materials

Interactions unraveled in a child-friendly room, where mother and child sat side-by-side on children stools, at a children's table; two smaller tables were placed next to the larger table, at participants' arms length, where toys were displayed. Across the room, two cameras were pointed at each participant, capturing audio and video. An additional webcam was centered in front of the table to monitor the interaction from the experimenter room. For higher fidelity of audio recording a lapel microphone was fixed to the table, on the mother side, that feed directly to a camera, and, on the children side, the microphone of a smartphone provided an additional audio source. All equipment was synchronized through a wireless synchronization system including a raspberry pi and a mobile app. The complete procedure consisted of several semi-structured play tasks, across two sessions, with different interactional partners. For the purpose of our research, we will focus on the more naturalistic task, that served as the baseline condition for further manipulations in the original research project. In this free-play task, mothers and their children played with the toys available on the side tables, for a total of 10 minutes. Mothers were instructed to play with their children as they would at home. Age appropriate toys were selected: coins set; dinosaur set; key; wooden puzzles set; cups set; plate; fork, knife, spoon set; construction blocks set; wooden blocks set; story book. Toys were randomly and evenly distributed between the side tables, in order that a similar number of toys were at grasp range from each participant. Before the beginning of the procedure, mothers were informed that they could stop the interaction at any time if they considered the child was visibly uncomfortable or uncooperative, but were encouraged to help their children engage in the task.

Coding of Gaps and Overlaps in Turn-Transitions

For the coding of mother and infant turn-transitions, vocalizations were manually segmented by marking their onset and offset, using the ELAN software (Wittenburg et al., 2006). Assistants were trained to segment vocalizations, the resulting segments were then automatically coded, following the on-off binary logic of the Automated Vocal Transaction Analyzer (AVTA; Cassotta et al., 1964) model to derive dyadic vocalization states to which turn ownership rules are applied, as suggested by the dialogical systems approach (Jaffe & Feldstein, 1970), to further derive dyadic states of turn-taking. To the resulting dyadic states of interruptive simultaneous speech (ISS) and switching pause (SP) correspond, respectively, overlaps and gaps. For further details on this approach, see Lourenço, Pereira, et al. (2021) where the same method is exposed.

Analysis of Turn-Transitions

We calculated the frequency of turn-transitions, separating overlaps and gaps by age groups (3 years-old, 4 years-old, 5 years-old) and the interactional partner (mother or child) producing the transition. We then calculated average and median floor-transfer offset (FTO), gaps and overlaps durations, by partner for the three age groups.

To assess the significance of the differences on turn-transition duration we used a model comparison approach and (generalized) linear mixed effects modeling. The unit of analysis was each individual turn-transition, i.e. each dyad contributed with multiple gaps and overlaps. For modeling the effects on FTO duration, we fitted a linear mixed model with a Normal distribution. For modeling gaps and overlaps separately, given that each distribution corresponds to cutting in two parts the unimodal distribution of the FTO, and no standard transformation made the two highly skewed distributions approximately Normally distributed, we fitted a generalized linear mixed model to a Gamma distribution instead. As fixed effects we included children's age as a continuous predictor and direction of turn-transition (mother, child) as a categorical predictor. The random effects included a random intercept per dyad.

Three best-fit models were selected using a model comparison approach, one with FTO duration, another with gap duration, and the last one with overlap duration, as dependent variable. The final model was selected by starting with a null model that included only the random effect and incrementally adding fixed effects and interaction terms. The effects that were significant according to a likelihood ratio were kept.

Models were fitted using the *lmer* and *glmer* functions of the *lme4* package (Bates et al., 2012) in R (R Development Core Team, 2012). Statistical inference was based on computing the estimated marginal means and corresponding 95% confidence interval. This was done using the R package emmeans (Lenth, 2021).

Results

Descriptive Statistics

Frequency

By applying the AVTA model to the segmented vocalizations a total of 7,208 turn-transitions were detected, of which 1.3% were excluded from the analysis, for being 3.5SD longer (4,954 ms) than the mean turn-transition duration of the initial sample, an artifact of the AVTA model's inability to distinguish gaps from the long pauses (time-outs) between bursts of conversation. From the remaining 7,113 turn-transitions, 3,570 were produced by the children, and 3,453 were produced by their mothers.

Table 4 displays the proportions of gaps and overlaps produced by children and mothers. For convenience, data was split into three age groups, corresponding to the natural age (in years) of the children: 3 years (n = 13), 4 years (n = 14) and 5 years (n = 17). Overall, results show that the proportion of gaps and overlaps in children is consistent across age groups, exhibiting similar proportions than those expected in adults.

Table 4

Proportion of Gaps and Overlaps by Partner and Age Group

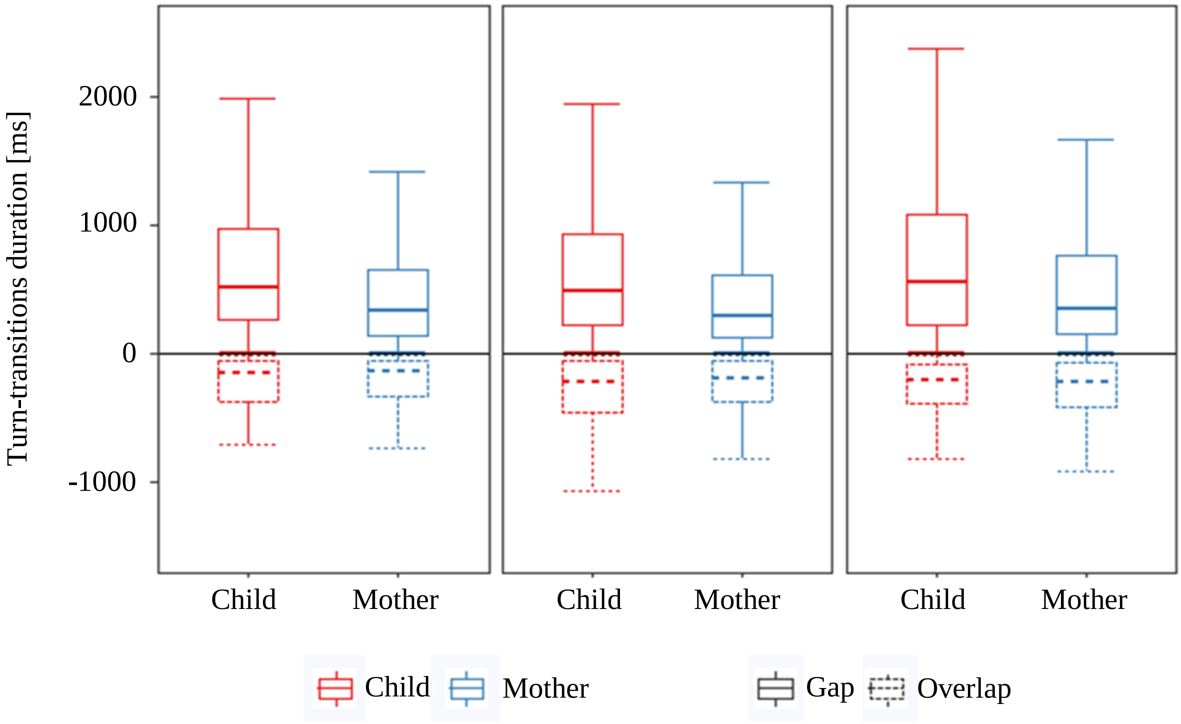
Partner	Transitions	3 Years	4 Years	5 Years	Total
Child	Gaps	88%	85%	85%	86%
	Overlaps	12%	15%	15%	14%
Mother	Gaps	77%	74%	82%	78%
	Overlaps	23%	26%	18%	22%

Duration

Gap durations were extracted from the switching pause (SP) dyadic state in the AVTA model, and assume positive values. Overlap durations were extracted from the interruptive simultaneous speech (ISS) dyadic state, but are presented as negative values for convenience, based on the floor-transfer offset convention. Average and median durations were calculated for turn-transition duration, per age group and direction of turn-transition, both as FTO, and as gaps and overlaps, independently. Figure 6 shows the distribution of gaps and overlaps duration.

Figure 6

Boxplot of the Distribution of Gaps and Overlaps Duration by Partner and Age Group



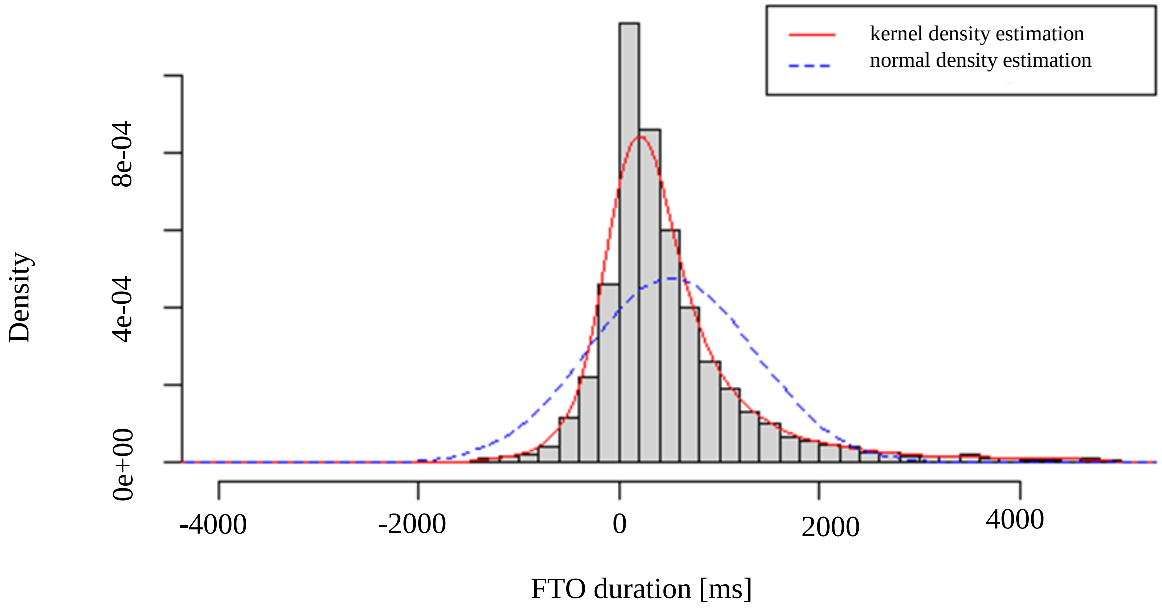
In the 3-years-old group, children's turn-transition durations averaged around 639 ms (*Mdn* = 440 ms), with mean durations of 764 ms (*Mdn* = 520 ms) for gaps, -280 ms (*Mdn* = -145 ms) for overlaps. Mother's turn-transition durations averaged around 377 ms (*Mdn* = 200 ms), with mean gap durations of 558 ms (*Mdn* = 315 ms) and mean overlap of -241 ms (*Mdn* = -140 ms).

In the 4-years-old group, turn-transitions initiated by children averaged around 579 ms (*Mdn* = 391 ms), with mean gap durations of 737 ms (*Mdn* = 485 ms), and overlap durations of -348 ms (*Mdn* = -220 ms). While turn-transitions initiated by mothers averaged around 345 ms (*Mdn* = -180 ms), with gaps of 563 ms (*Mdn* = 300 ms) and overlaps of -292 ms (*Mdn* = -190 ms).

In the 5-years-old group, children's turn-transition durations averaged around 672 ms (*Mdn* = 440 ms), with mean durations of 844 ms (*Mdn* = 560 ms) for gaps, and of -286 ms (*Mdn* = -200 ms) for overlaps. Turn-transitions initiated by mothers averaged around 468 ms (*Mdn* = 296 ms), with mean gap durations of 639 ms (*Mdn* = 349 ms), and mean overlap durations of -296 ms (*Mdn* = -215 ms). Finally, the dyadic frequency distribution of turn-transition duration, measured as FTO, illustrated in Figure 7, shows a slightly skewed and unimodal distribution, similar to what is expected in adults (Levinson & Torreira, 2015; Stivers et al, 2009).

Figure 7

Histogram of Density Estimation for the Probability Distribution of Turn-transition Duration (Measured as FTO)



Density functions estimations, show that the distribution of turn-transition approximates to a Normal distribution. The same is not expected if the distribution is split between negative (overlaps) and positive (gaps) values, and that should be taken in consideration when modeling gaps and overlaps, independently.

Modeling of Turn-Transition Duration

In line with the delineated research objectives, turn-transition duration was modeled first as floor-transfer offset (FTO), aggregating gaps and overlaps in a single metric that integrates both dimensions of the phenomenon. After, gaps and overlaps were considered separately.

Floor-Transfer Offset (FTO)

To understand the predictor effect of child's age and direction of turn-transition (child, mother) in turn-transition duration as a whole, we used the floor-transfer offset (FTO) as a dependent variable, which combines gaps and overlaps in a single measure where overlaps assume negative values and gaps positive values. We used a model comparison approach. Given the approximation of FTO distribution to a normal distribution (see Figure 7), linear mixed models were used.

A base model, with FTO duration as a dependent variable and a random intercept per dyad, as a random effect, was built. We then compared it to models including age, direction and the interaction between both factors, as fixed effects. Only the addition of direction of turn-transition significantly improved the model, $\chi^2(1) = 142.63$, $p = < .001$. Although age did not significantly improved the model ($\chi^2(1) = 1.175$, $p = 0.278$) it was added as a factor to the final model, in order to test the developmental hypotheses. Table 5 shows the results for the final model (*log-likelihood* = -8,631.2, N = 7,113).

Table 5

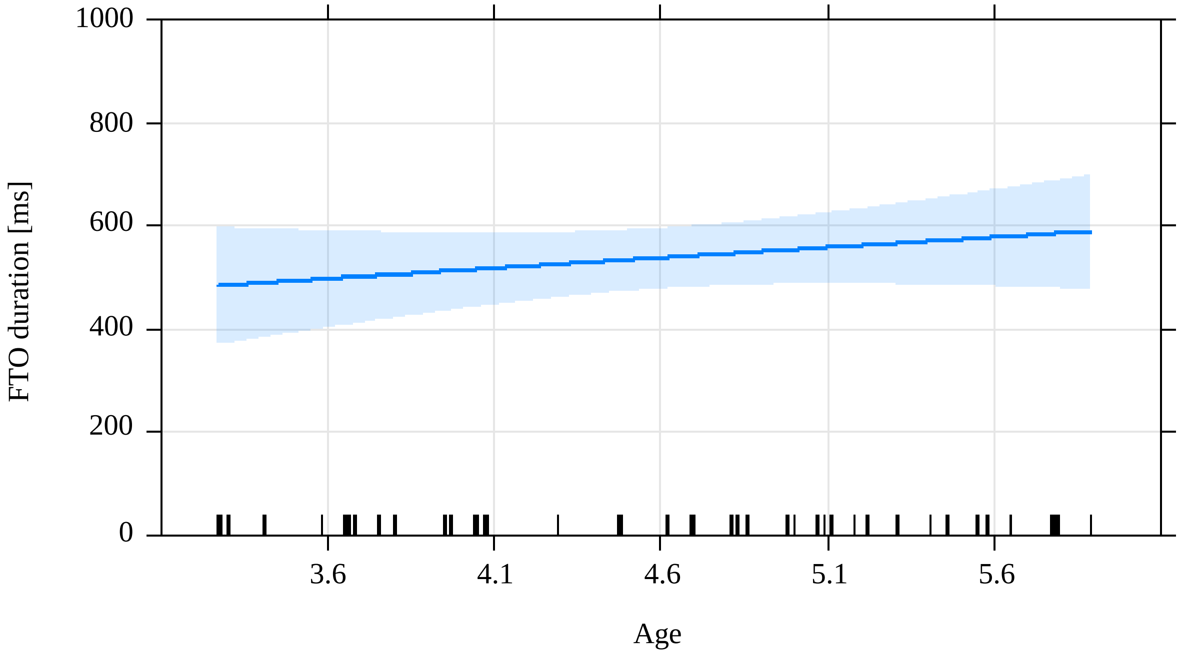
Parameter Estimates, Confidence Intervals, and Significance Values for the Fixed Effects in the FTO Duration Model

Predictors	Estimates	CI	*P*
Intercept	0.54	0.48 – 0.59	<0.001
Age	0.04	-0.03 – 0.11	0.284
Direction of Turn-Transition (child)	0.12	0.10 – 0.13	<0.001

There was no significant main effect of age in dyadic FTO duration (β = 1.08, CI = 0.98-1.19, p = 0.101), with Figure 8 showing only a slight increase in FTO duration throughout age points.

Figure 8

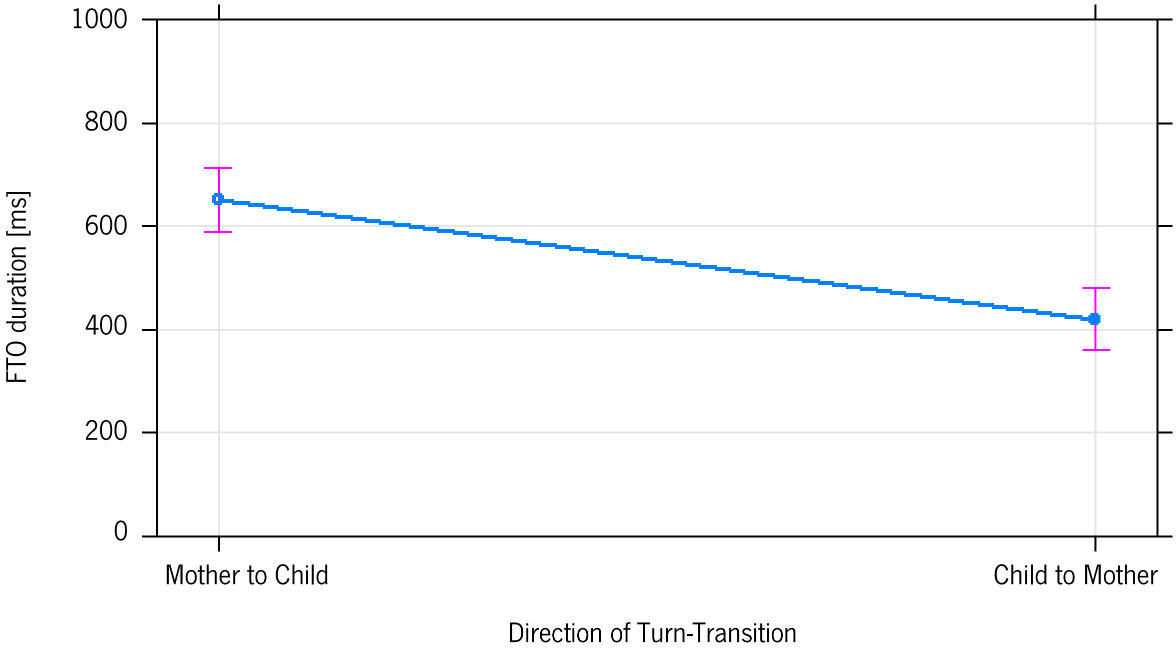
Predictor Effect Plot for Age in the FTO Duration Model



On the other hand, a significant main effect was found for the direction of turn-transition (β = 0.12, *CI* = 0.10-0.13, $p < 0.001$). Figure 9 plots the estimated marginal means and confidence intervals for the predictor effect of direction, showing that children initiated turn-transitions are, on average, longer (*M* = 650 ms) than the gaps initiated by their mothers (*M* = 420 ms).

Figure 9

Predictor Effect plot for Direction in the FTO Duration Model



Given the significant effect of the direction of turn-transition, additional explorations to understand the possible effect of the child's age in FTO duration were conducted, using the same model comparison approach, but analyzing children and mothers turn-transitions independently. None of the explorations provided evidence of any significant effect of age in children or mother FTO duration.

Gaps

To understand the effects of child's age and direction of turn-transition (child, mother) specifically on gap duration, we again used a model comparison approach, but this time utilizing general linear mixed models with a Gamma distribution for the dependent variable, to account for the distributional split between gaps and overlaps.

We first built a model with gap duration as a dependent variable and a random intercept per dyad, as a random effect. We compared this base model to models including age, direction and the interaction between both factors, as fixed effects. Again, only the addition of direction of turn-transition significantly improved the model, $\chi^2(1) = 125.76$, $p < .001$. As previously, although age did not significantly improved the model ($\chi^2(1) = 2.641$, $p = 0.104$) we added it to the final model to test for developmental changes. Table 6 shows the results for the final model (*log-likelihood* = -3,529.1, *N* = 5831).

Table 6

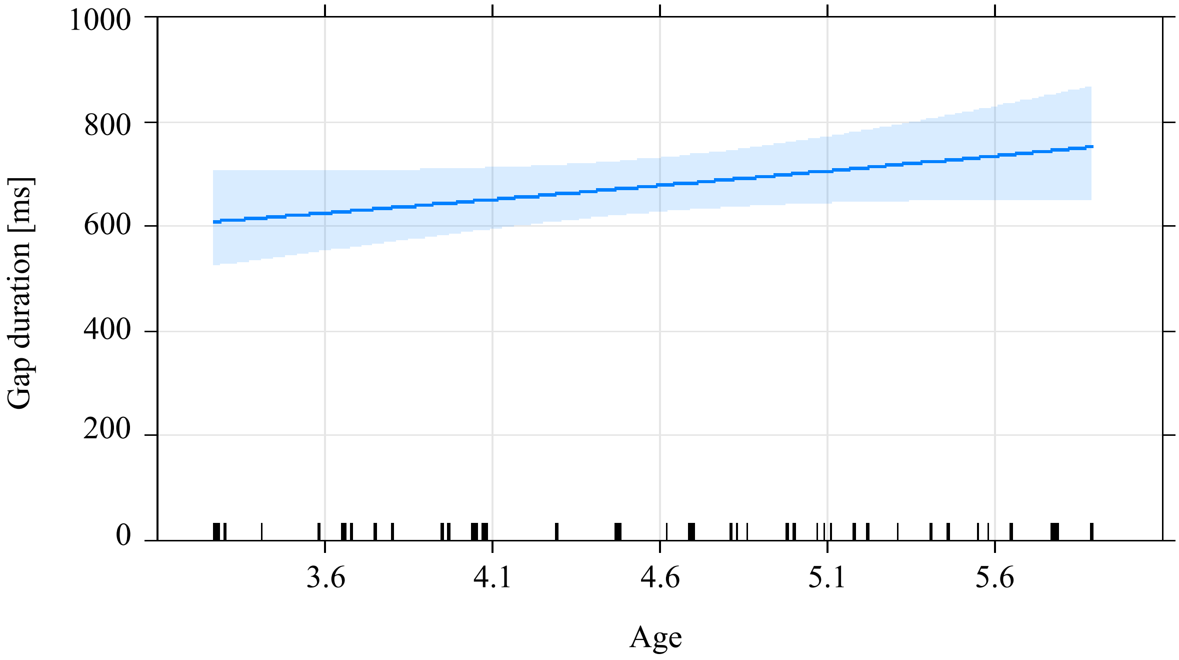
Parameter Estimates, Confidence Intervals, and Significance Values for the Fixed Effects in the Gap Duration Model

Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.67	0.62 – 0.72	<0.001
Age	1.08	0.98 – 1.19	0.099
Direction of Turn Transition (Child)	1.16	1.13 – 1.19	<0.001

Once again, no significant main effect of age in dyadic gap duration (β = 1.08, *CI*= 0.98-1.19, *p* = 0.101) was found. Nevertheless, Figure 10 still shows a more accentuated increase tendency over time in dyadic gap duration, than with FTO duration.

Figure 10

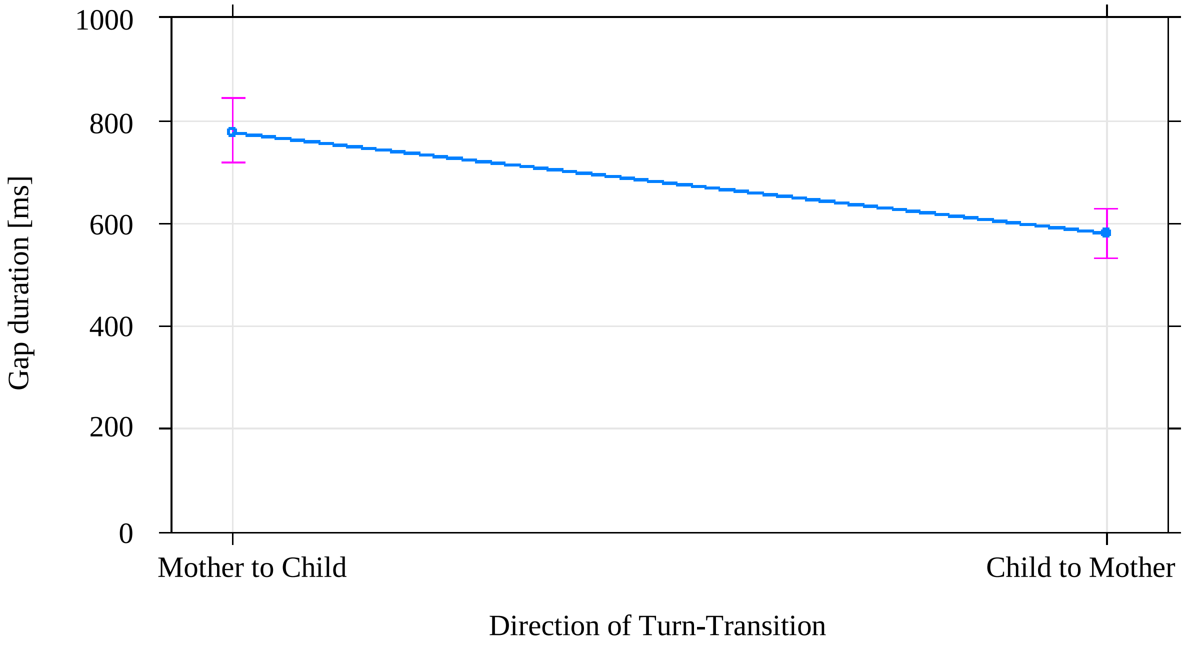
Predictor Effect Plot for Age in Gap Duration



Similarly to FTO, a significant main effect was found in gap duration for the direction of turn-transition (β = 1.16, *CI* = 1.13-1.19, *p* < 0.001). Figure 11 plots the estimated marginal means and confidence intervals for the predictor effect of direction, showing that, as with FTO duration, children-initiated gaps are also, on average, longer (*M* = 778 ms) than the gaps initiated by their mothers (*M* = 579 ms).

Figure 11

Predictor Effect Plot for Direction in the Gap Duration Model



Given the significant effect of direction of turn-transition, we similarly conducted additional explorations to understand the possible effect of the child's age in gap duration, by analyzing children and mothers turn-transitions independently. Again, none of the explorations provided evidence of any significant effect of age in gap duration.

Overlaps

Finally, to understand the predictor effect of child's age and direction of turn-transition (child to mother, mother to child) in overlap duration, we follow the same model comparison approach, utilizing general linear mixed models with a Gamma distribution. Differently from FTO convention, overlap duration was modeled as a quantity assuming positive values.

As previously, we first built a base model, with overlap duration as a dependent variable and a random intercept per dyad, as a random effect. We then compared this base model to models including age, direction, and the interaction between both factors, as fixed effects. None of the factors improved the base model. Although neither age ($\chi^2(1) = 0.516$ $p = 0.472$), nor direction of turn-transition ($\chi2(1) = 1.967$, $p = 0.373$) significantly improved the model, we added both to the final model, in order to compare it with FTO and gap results. Table 7 shows the results for the final model (*log-likelihood* = 330.60, *N* = 1,282).

Table 7

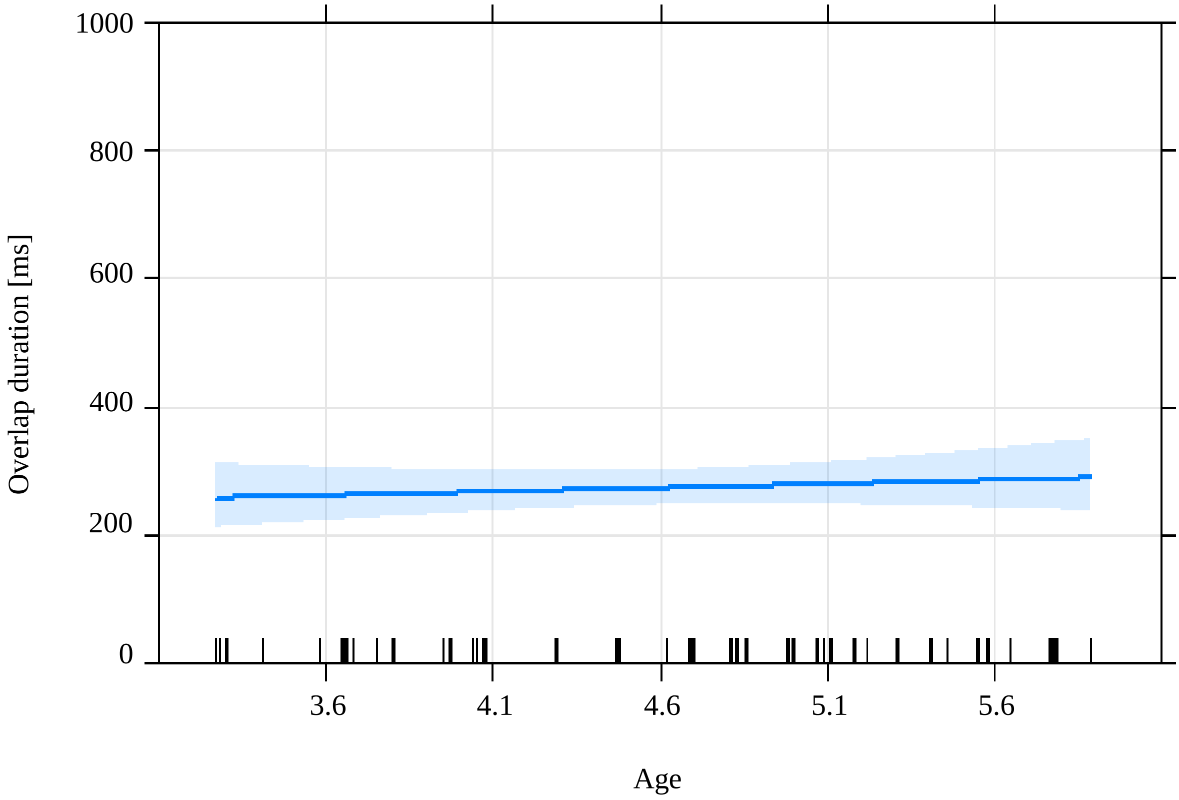
Parameter Estimates, Confidence Intervals, and Significance Values for the Fixed Effects in the Overlap Duration Model

Predictors	Estimates	CI	*P*
Intercept	0.27	0.25 – 0.30	<0.001
Age	1.04	0.92 – 1.19	0.469
Direction of Turn-Transition (child)	1.04	1.13 – 1.19	0.229

No significant main effect of age in dyadic overlap duration ($\beta = 1.04$, *CI* = 0.92-1.19, $p = 0.469$) was found, visible in Figure 12 that shows a much more flattened evolution from age point to age point than with gaps, or even FTO durations.

Figure 12

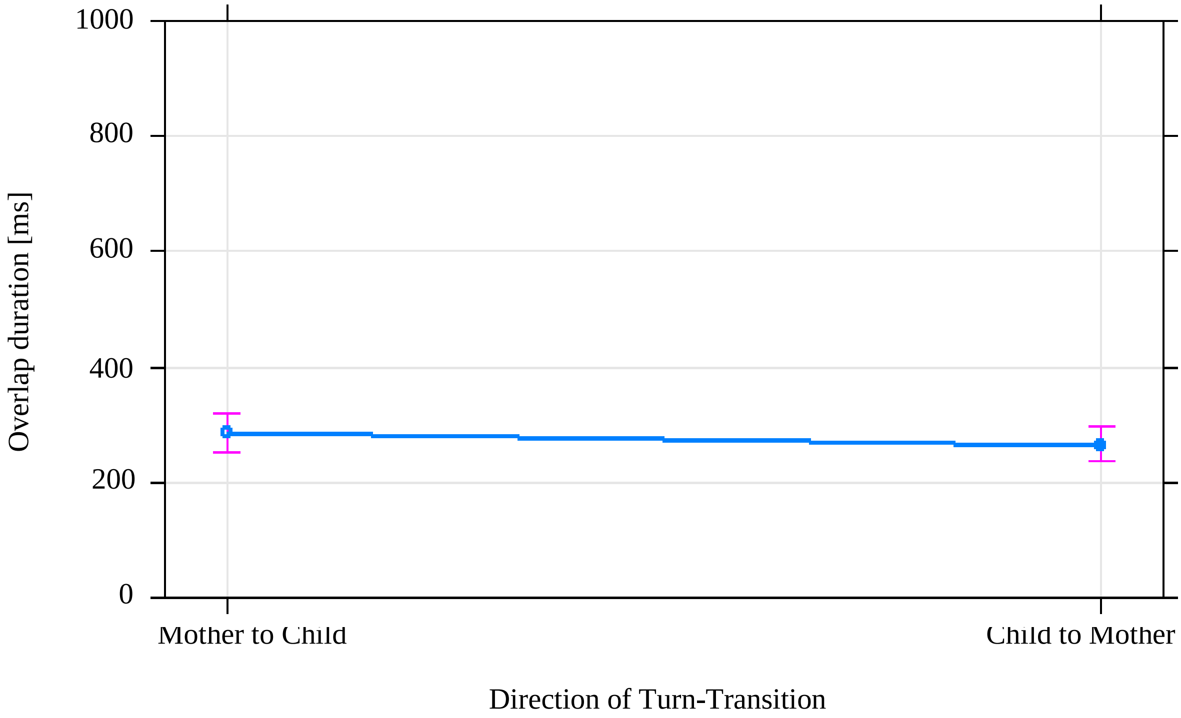
Predictor Effect Plot for Age in the Overlap Duration Model



Likewise, there was no significant main effect of the direction of turn-transition in overlap duration (β = 1.04, *CI*= 1.13-1.19, *p* = 0.229), as illustrated by Figure 13.

Figure 13

Predictor Effect Plot for Direction in the Overlap Duration Model



Discussion

Our study provides a microanalytic perspective on the development of turn-taking in preschoolers. Most recent developmental studies of turn-taking (Casillas et al., 2016; Hilbrink et al., 2015; Stivers et al., 2018) have a markedly psycholinguistic approach that builds on the interaction engine hypothesis (Levinson, 2006, 2019). That work has focused on the linguistic processing implications in turn-transition timing of the integration between a more foundational interaction system, and the emerging language system. When comparing infancy and childhood research, important methodological differences can be found. While in infant studies all preverbal vocalizations can be object of analysis, in studies with children there is a shift towards restricting the analysis to semantically contingent transitions, most commonly question-response pairs. Likewise, while in infancy the differential trajectory of gap and overlap duration is emphasized, in childhood, aggregated measures of gap and overlap are preferred, which may obscure the relative contribution of each dimension to turn-transition timing.

Our approach differentiates from these methodological options, first, by focusing on the whole spectrum of vocalizations, privileging temporal contingency to semantic contingency. Second, by presenting both aggregated and independent analysis of gap and overlap duration.

Overall, our results suggest that, between 3;3 and 5;10 years, there is no evidence of an age effect (dyadic or partner-dependent) on turn-transition duration. Neither when gaps and overlaps are considered independently or as dimensions of the same phenomenon, when using the floor-transfer offset. On the other hand, there is a noticeable difference in the estimated duration of children's turn-transitions (650 ms), when compared to their mothers (420 ms), that would suggest that preschoolers are still not converging to adult standards on turn-transition timing.

An independent analysis of both types of turn transitions – gaps and overlaps – though, shows a more complex picture of the phenomenon. Similarly to the FTO results, children's estimated gap duration (778 ms) maintain a noticeably from their mothers (579 ms). Meanwhile, children's overlap durations did not significantly differ from the overlap durations of their mothers.

These results reinforce the evidence that gaps and overlaps appear to have a different developmental trajectory, and may contribute differently for the timing of children's turn-transitions (Hilbrink et al., 2015; Lourenço, Pereira, et al., 2021).

We will proceed by comparing our results with other studies in children to understand how our microanalytic approach to what is defined as turn-transition (all temporal contingent vocalizations) compares to the psycholinguistic approach.

Comparison with Children Research

Compared to the median response latencies reported for all six combinations question-answers in Casillas et al. (2016), for all ages (575 ms), and specifically at 3;0-3;1 (571 ms) and 3;3-3;5 (523 ms) years old, our median FTO results for the 3-years-old group are just slightly shorter (440 ms). In fact, our result is even closer to the timing of the 3;3-3;5 years old, when only yes/no questions are considered (465 ms).

When comparing the median response latencies for all pair-adjacent turn-transitions in Stivers et al. (2018), for the 4-5 age group (500 ms), again our median FTO results for the 4-years-old (391 ms) and 5-years-old group (440 ms) show slightly shorter durations, of similar magnitude (~100 ms).

The magnitude and direction of these differences is most interesting, given the methodological differences in what was considered as turn-transition by the studies using a psycholinguistic approach, that only examine semantically contingent pair-adjacent vocalizations (Casillas et al., 2016; Stivers et al., 2018), and our study that considered all temporal contingent vocalizations. This suggests that our approach may be including shorter turn-transitions at different levels than those analyzed in the psycholinguistic studies; and may indeed be a better general measure of turn-transition duration in childhood, which can be used as a reference and a base line for studies that focus on specific types of turn-transitions (e.g., different levels of linguistic complexity).

We will proceed to compare the results in our study for gap and overlap durations, with those reported in the study of infants (Hilbrink et al., 2015; Lourenço, Pereira, et al., 2021) to understand how they might differ, as an indication of the developmental trajectory of each dimension.

Comparison with Infant Research

Regarding gaps, the median durations reported in Hilbrink et al. (2015) for the two furthest age points are of around 975 ms, at 12 months, and around 700 ms, at 18 months. Median gap durations reported in Lourenço, Pereira, et al. (2021) at 12 months, are of 570 ms, when toys are available, and 279 ms, when toys are removed from the interaction. Comparing both studies to the median results in our study for the 3-years-old group (520 ms), produces alternative explanations for the trajectory of gap duration between late infancy and early childhood.

If we take the results in Hilbrink et al. (2015) as reference it would appear that gap duration may be getting shorter throughout that period. This would be coherent with the authors' suggestion that the integration between the interaction system and the language systems begins at an earlier stage (9 months) and develops throughout infancy. This however is not consistent with the results of a recent meta-analysis (Nguyen et al., 2022), which suggests an ascending trajectory in gap duration, at least up until 40 months. Nor with the studies of response latencies throughout this period, and beyond, that suggest that turn-taking timing has a slow progression towards the minimal-gap minimal-overlap standard in adulthood (Casillas et al., 2016; Stivers et al., 2018).

If we consider the results in Lourenço, Pereira, et al. (2021), for the free-play with toys task (570 ms), as a reference, it would suggest that gap durations may maintain similar durations between late infancy and early childhood. Still, the authors have demonstrated that by removing objects from the interaction, gap durations could be much shorter. If we then take the results of the free-play without toys task (279 ms), it would be more appropriate to consider that gap durations may be increasing. This last interpretation would also be the coherent with the trajectory proposed by Nguyen et al. (2022) and the most reflective of the spectrum of developments in language processing throughout this period (Casillas et al., 2016).

Concerning overlaps, median durations reported in Hilbrink et al. (2015), are relatively stable from 9 to 18 months (~ 675 ms – values are positive because FTO was not used). And, in Lourenço, Pereira, et al. (2021) median overlap durations, at 12 months, are of -246 ms, with toys, and -360 ms, without toys.

Again, there are implications between which values we use as reference, but this time only on the magnitude of the difference, since all results points towards a descending trajectory, towards adult-like durations, that only slightly varies over time– 3 year-old group (-145 ms), 4 year-old group (-220 ms), and 5 year-old group (-180 ms). These minor differences, plus our supplementary analysis that found no significant differences between children’s and their mothers overlaps, is a strong indication that the convergence towards minimal-overlap may be locked, somewhen between late infancy and early childhood.

Conclusion

Taking it all together, we interpret the results from the present study and other developmental studies of turn-taking as a clear evidence that gap and overlap durations continue to have different developmental trajectories beyond infancy. Overlap durations are getting shorter, and converge to adult standards sometime between late infancy and early childhood. Gap durations may increase throughout early childhood (Nguyen et al., 2022), but progress in a more non-linear fashion: maintaining similarly longer durations, with some variation between different levels of linguistic complexity (Casillas et al., 2016), and a slow progression towards minimal-gap – that may prolong in time, even beyond the 8 years mark (Lindsay et al., 2019; Nguyen et al., 2022; Stivers et al., 2018).

Furthermore, if we take these differences in the developmental trajectory of gaps and overlaps into account, we can, at least from 3 years old onwards attribute most of the differences between children and adult turn-transition timing to differences in gap duration.

Nevertheless, the methodological differences between studies should be considered carefully when comparing the results of other studies to our own, to avoid erroneous extrapolation. To sum up, our study demonstrates that (1) measuring turn-transitions in childhood, by considering all temporally contingent vocalizations can provide, at least, similar results to those when only chosen semantically contingent turn-transitions are analyzed – and perhaps be a better general measure to be used as reference; and that (2) using complementary metrics of turn-transition duration, such as the floor-transfer offset (FTO), gaps and overlaps, can help us understand the contribution of each dimension to the timing of turn-transitions. Additionally, we presented evidence that (3) gaps and overlaps continue to have different developmental trajectories throughout childhood; and that (4) there are strong indicators that, at least, by 3 years old, overlap duration has converged to the minimal-overlap standard of adulthood.

We believe that these results have relevant methodological implications for future research into the development of turn-taking, and improve our understanding of the developmental trajectory of turn-transition timing.

References

Bates, D., Maechler, M., & Dai, B. (2012). *Lme4: Linear mixed effects models using s4 classes.*

Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, *263*(1), 101–113. https://doi.org/10.1111/j.1749-6632.1975.tb41575.x

Beebe, B., Alson, D., Jaffe, J., Feldstein, S., & Crown, C. (1988). Vocal congruence in mother-infant play. *Journal of Psycholinguistic Research*, *17*(3), 245–259. https://doi.org/10.1007/BF01686358

Brazelton, T. B., Koslwski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). Willey-Interscience.

Casillas, M. (2014). Turn-taking. In D. Matthews (Ed.), *Pragmatic development in first language acquistion* (pp. 53–70). Benjamins. https://doi.org/10.1075/tilar.10.04cas

Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn-taking, timing, and planning in early language acquisition. *Journal of Child Language*, *43*(6), 1310–1337. https://doi.org/10.1017/S0305000915000689

Casillas, M., & Frank, M. (2013). The development of predictive processes in children's discourse understanding. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 299–304. http://langcog.stanford.edu/papers/CF-cogsci2013.pdf

Cassotta, L., Feldstein, S., & Jaffe, J. (1964). AVTA: A device for automatic vocal transaction analysis 1. *Journal of the Experimental Analysis of Behavior*, *7*(1), 99–104. https://doi.org/10.1901/jeab.1964.7-99

Elias, G., Hayes, A., & Broerse, J. (1986). Maternal control of co-vocalization and inter-speaker silences in mother-ineant vocal engagements. *Journal of Child Psychology and Psychiatry*, *27*(3), 409–415. https://doi.org/10.1111/j.1469-7610.1986.tb01842.x

Ervin-Tripp, S. (1979). Children's verbal turn-taking. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 391–414). Academic Press.

Fogel, A. (1977). Temporal organization in mother-infant face-to-face interaction. *Studies in Mother-Infant Interaction*, *January 1977*, 119–152.

Garvey, C., & Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse Processes*, *4*(1), 27–57. https://doi.org/10.1080/01638538109544505

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology*, *6*(September), 1–12. https://doi.org/10.3389/fpsyg.2015.01492

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), i–viii, 1–132. http://www.ncbi.nlm.nih.gov/pubmed/11428150

Jasnow, M. D., & Feldstein, S. (1986). Adult-like temporal characteristics of mother-infant vocal interactions. *Child Development*, *57*(3), 754–761. http://www.jstor.org/stable/1130352

Kaye, K., & Fogel, A. (1980). The temporal structure of face-to-face communication between mothers and infants. *Developmental Psychology*, *16*(5), 454–464. https://doi.org/10.1037/0012-1649.16.5.454

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means. R package version 1.5.5-1.* https://cran.r-project.org/package=emmeans

Levinson, S. C. (2006). On the human interactional engine. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality. Culture, Cognition and Interaction* (pp. 39–69). Berg.

Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 103–130). Willey-Blackwell.

Levinson, S. C. (2019). Interactional foundations of language: The interaction engine hypothesis. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 189–200). MIT Press.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*(JUN), 1–17. https://doi.org/10.3389/fpsyg.2015.00731

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychological Science*, *30*(4), 504–515. https://doi.org/10.1177/0956797618822802

Lourenço, V., Coutinho, J., & Pereira, A. F. (2021). Advances in microanalysis: Magnifying the social microscope on mother-infant interactions. *Infant Behavior and Development*, *64*(April), 101571. https://doi.org/10.1016/j.infbeh.2021.101571

Lourenço, V., Pereira, A. F., Sampaio, A., & Coutinho, J. (2021). Turn-Taking in object-oriented and face-to-face interactions: A longitudinal study at 7 and 12 months. *Psychology & Neuroscience*. Advance online publication. https://doi.org/10.1037/pne0000276

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social Perception in Infants* (pp. 177–197). Ablex Publishers.

Nguyen, V., Versyp, O., Cox, C., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the development of turn taking in adult–child vocal interactions. *Child Development*, 1–20. https://doi.org/10.1111/cdev.13754

R Development Core Team. (2012). *R: A language and environment for statistical computing.* r foundation for statistical. http://www.r-project.org/

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Stern, D. N. (1971). A micro-analysis of mother-infant interaction: Behavior regulating social contact between a mother and her 3 1/2-month-old twins. *Journal of the American Academy of Child Psychiatry*, *10*(3), 501–517. https://doi.org/10.1016/S0002-7138(09)61752-0

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stivers, T., Sidnell, J., & Bergen, C. (2018). Children's responses to questions in peer interaction: A window into the ontogenesis of interactional competence. *Journal of Pragmatics*, *124*, 14–30. https://doi.org/10.1016/j.pragma.2017.11.013

Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies*, *7*(2), 251–271. https://doi.org/10.1075/is.7.2.08str

Thorgrimsson, G. B. (2014). *Infants' understanding of communication as participants and observers*. [Unpublished doctoral dissertation]. Radbound University Nijmegen.

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Tronick, E., Als, H., & Brazelton, T. B. (1977). Mutuality in mother-infant interaction. *Journal of Communication*, *27*(2), 74–79. https://doi.org/10.1111/j.1460-2466.1977.tb01829.x

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of the 5th International Conference on Language Resources and Evaluation, LREC 2006*, *August*, 1556–1559.

CHAPTER IV – STUDY 3[3]

---

[3] A version of this chapter was published in the Elsevier peer-reviewed journal *Infant Behavior & Development*, for the special issue in Methodological Advances in the Characterization and Understanding of Caregiver-Infant Interactions:

Lourenço, V., Coutinho, J., & Pereira, A. F. (2021). Advances in microanalysis: Magnifying the social microscope on mother-infant interactions. *Infant Behavior and Development*, *64*(April), 101571. https://doi.org/10.1016/j.infbeh.2021.101571

Abstract

Microanalysis is a method for recording and coding interactional behavior. It has been often compared to a social microscope, for its power in detailing the second-by-second dynamics of social interaction. Microanalysis has deep multidisciplinary foundations, that privilege the description of interactions as they naturally occur, with the purpose of understanding the relations between multiple and simultaneous streams of behaviors. In developmental science, microanalysis has uncovered structural and temporal elements in mother-infant interactions, improving our understanding of the effects of mother-infant interpersonal adaptation in the infant's cognitive and social-emotional development.

Detailed manual coding is time intensive and resource demanding, imposing restrictions to sample size, and the ability to analyze multiple behavioral modalities. Moreover, recent increases in the density of multivariate data require different tools. We review present-day techniques that tackle those challenges: (1) sensing techniques for motion tracking and physiological recording; (2) exploratory techniques for detecting patterns from high-density data; and (3) inferential and modeling techniques for understanding contingencies between interactional time series.

Two illustrations, from recent developmental research, reveal the power of bringing a new lens to our social microscope: (1) egocentric vision, the use of head mounted cameras and eye-trackers in capturing the infant's first-person perspective of a social exchange; and (2) daily activity sensing, wearable multimodal sensing that brought mother-infant interaction research to the environments where it naturally unfolds.

Keywords: mother-infant interaction, microanalysis, interpersonal adaptation, interactive contingency, time series models

Advances in Microanalysis: Magnifying the Social Microscope on Caregiver-Infant Interactions

The quality of the infant's early relationships has a meaningful impact in infant's cognitive, emotional, and social development (Leclère et al., 2014). Mother-infant interactional behaviors unravel at different timescales across behavioral channels; how these overt behaviors organize and unfold over time can inform us about psychological processes, helping us understand normal and abnormal development.

From the groundbreaking studies, to recent work, studying the temporal structure of interactive behavior has been recognized as a major challenge (de Barbaro et al., 2013; Fogel, 1977; Xu et al., 2020). In adults, social psychology studies have documented a large range of interpersonal adaptation patterns (Burgoon et al., 1995), and emphasized the temporal sequence of moment-to-moment behaviors as a core dimension of interactions (Bakeman & Gottman, 1997; Bakeman & Quera, 2011; Gottman, 1990). Social partners can flexibly rearrange the multimodal flow of an interaction, changing behavior contingently to each other, by responding in a similar direction, or with a dissimilar behavior; this can happen in both partners or only unilaterally. In general, interactive contingency effects occur within and across behavioral channels, at times as a multimodal package that congruently changes in the same direction, at other times with incongruent changes across modalities (Burgoon et al., 1995; Cappella, 1981). Interactions also exhibit rhythmicity qualities that are irregular or non-periodic (Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001). Even the immature infant can engage with an adult at this level of sophistication (Beebe et al., 2010).

This poses multiple difficulties to understanding the real-time dynamics of any interaction, not just the mother-infant dyad. In a multimodal context, how to record and code several behavioral channels? If the stream of joint behaviors occurs in parallel and in both partners, how to explore the large space of potential relations between them? How to adequately model multivariate time-series data of this nature? The statistical properties of social exchanges, including mother-infant interactions, also pose additional complexities in statistical inference and modelling. One is the strength of interactive contingency: this quantity is not constant – e.g. the magnitude of the statistical association, at some temporal lag, between two specific behavioral streams; interactive contingency is subtly context sensitive and can change within the same interaction. In time series terminology, the data exhibits non-stationarity. This is simple to point out, for instance, in a face to face interaction at 4 months: mother and infant are not in a constant level of activity and engagement; the interaction waxes and wanes, with sustained periods of stability that are punctuated by moments of low activity, silence, or disruption, e.g. see the "disruption and repair" pattern, in studies of infant attachment (Beebe & Lachmann, 2013). Outside the confines of the laboratory setting, and the structured social interaction task, recent studies have found that infants experience social interactions in bursts of activity, punctuated by long periods of silence (Tamis-LeMonda et al., 2017). The amazing flexibility of social exchanges entails stationarity at short time scales, at best, and this has major implications for all available statistical techniques.

One of the earlier approaches in developmental science to tackle this challenge was microanalysis (Bateson, M. C., 1971; Brazelton, Koslwski, & Main, 1974; Condon & Sander, 1974; Stern, 1971), a method that emerged from interdisciplinary efforts in understanding human communication (McQuown, 1971). In its strict sense of a technique, microanalysis is a method for detailed observation and coding of recorded interactional behavior, typically on a split-second time scale. More broadly, it is an approach with theoretical and empirical implications for research – and historically, has exploited technological evolution in order to enhance our analytical ability. Microanalytic studies typically examine the real-time structure of social interaction as it occurs in intact dyadic situations, either in a controlled recording environment in the laboratory (Brazelton et al., 1975; Tronick et al., 1978) or in the infant's socio-cultural context (Bateson, M. C., 1975; Stern, 1971).

A useful metaphor is of a social microscope, that applies audio and video recording technology (and others), along with frame-by-frame reproducing techniques, to the multimodal flow of social interaction (Beebe, 2014). Just as the use of the first microscope by Antoni van Leeuwenhoek led to the discovery of a new world of microorganisms, invisible to the human eye, the use of cameras and sound recording equipment brought a paradigmatic change in the ability to examine interaction in detail. As Stern would later summarize it, "*When you have the [..] opportunity to be among the first people to see a new world, many of its surprising features are striking enough that they force you to reevaluate your preconceptions. You quickly grasp a new perspective and new realities (...)*" (Stern, 2002, p. 2). Today, a multitude of instruments and techniques (some only recently introduced) function as novel types of lenses on our social microscope. We will argue that, as before, this is changing our understanding of mother-infant interactions. Recent work in egocentric vision, that examines the natural statistics of visual experience from the infant's first-person perspective (see Smith, Jayaraman, Clerkin, & Yu, 2018, for a recent review), in corpus analysis of long recordings of infants carrying wearable sensors in their daily lives (Casillas & Cristia, 2019; Cychosz et al., 2020), and in highly controlled multimodal recordings of mother-infant interactions (Schroer et al., 2019), to list a few, have provided new hypotheses, derived from the opportunity of being the first to look inside the infant's social and cognitive ecology.

Our goal in this work is to present microanalysis' original motivations and findings, in the 1960's and 70's, and relate them to the potentialities of present-day techniques. Advances in sensors and data analysis techniques have enabled a major increase in the speed at which a microanalytic study can be concluded (Smith et al., 2018), and pushed towards examining social interaction in context (Cychosz et al., 2020; de Barbaro, 2019), as it occurs, a form of "natural history" of social behavior (Bateson, G., 1971).

The paper is structured in four main parts, followed by a discussion. First we will examine the emergence of the microanalytic method in its broad historical and theoretical context, along with the fundamental aspects that characterize it as a method: its object of analysis, the context of application, and the level of detail. We follow with a review of the seminal studies using microanalysis in infancy. In the third section we survey a subset of current techniques for recording, measuring, processing, analyzing, and modeling nonverbal interactional behavior that can alleviate some of the strenuous demands of classical microanalytic observational and coding methods. Our survey is organized into three broad categories: sensing, exploration, and modelling. In sensing technologies we report on some innovations in motion-tracking, multimodal capture of behavior, and measurement of physiological and neural responses. These novel technological options increase data resolution, are easier to use with less barriers to adoption, and also partially automate data segmentation and coding. In the second category, we present exploratory techniques to detect patterns within streams of multivariate data and support hypothesis generation. Finally, we will review a subset of modeling and hypothesis testing techniques to address the intrinsic complexity of interactive contingency. Our historical account of early microanalytic research, and review of contemporary methods for measuring the microdynamics of caregiver-infant interactions, will be necessarily selective, to fit within space constraints. Our criteria was to concentrate on the main aspects of the seminal studies, reflect on the similarities with current work, and examine techniques with clear possibility of new directions.

As a demonstration of the potential of microanalysis for changing our understanding of infancy, we will present two recent case studies. First is egocentric vision, a new field in developmental research concerned with the statistics of visual experience in the infant's first-person view (using head-mounted cameras and eye trackers); this area has used a literal new lens on infancy and brought a new perspective to the study of caregiver-infant interactions (Franchak et al., 2010; Smith et al., 2018; Yoshida & Smith, 2008). A second use case focuses on the potential of daily life activity sensing for increasing the ecological validity of developmental research (de Barbaro, 2019), echoing the original approach of ecological momentary assessment, and long-standing arguments of socio-cultural researchers of studying social interaction in its intact ecology. In the discussion we address the advantages and potentialities of bringing the new techniques to microanalytic research, as well as the limitations to their application and potential solutions.

A final remark on terminology. The intricacies of social interactions, specifically the coordinated aspects of the multimodal envelope of activity, do not have an agreed upon unique term: synchrony, coordination, interactive contingency, interpersonal adaptation, mutual regulation, linkage, coupling, etc. In a corpus study of the social coordination literature, across many fields, Paxton (2015) provides evidence for this heterogeneous conceptual landscape; as a necessary simplification, we consider all these terms as equivalent, as they all speak to social coordination in a broad sense (Paxton, 2015).

Foundations of Microanalysis

We begin by considering some of the contextual and theoretical influences that prompted the emergence of microanalysis, and reflect on the fundamental elements of microanalysis as a method. For a more in-depth account of historical events and ideas, see Bull (2002) or Leeds-Hurwitz (1987). Microanalysis emerged in a time of singular interdisciplinary convergence. On the aftermath of World War II, the Macy Conferences on cybernetics (1946-1953) and group processes (1954-1960) promoted the encounter and debate of the most exciting ideas in behavioral and social science that would later influence the microanalytic approach (Bull, 2002; Leeds-Hurwitz, 1987). They also served as a prelude to one of the most ambitious and unknown projects in behavioral science (Hutchins, 2010): the Natural History of an Interview (NHI; McQuown, 1971).

The Natural History of an Interview Project

The NHI project began by 1955, led by psychiatrist Frieda Fromm-Reichman, and main contributions of anthropologist Gregory Bateson, kineticist Ray Birdwhistell, psychiatrist Henry Brosin, linguists Charles Hockett, Norman McQuown, Henry Smith, and paralinguist George Trager (McQuown, 1971). Other major figures had their contribution in specific phases of the project – Alfred L. Kroeber, David M. Shneider, Erick Erickson; including soon to be protagonists of microanalytic research: Starkey Duncan, Jr., William S. Condon, Adam Kendon and Albert L. Scheflen, who were involved as members of the teams responsible for the analysis.

The objective was to make a detailed analysis of speech and body motion, using a film-recording of a single social interaction, involving an interviewer (Bateson) and his former psychiatric patient ("Doris"), at her home, and in the presence of her husband and son, and a cameraman (McQuown, 1971). The project was framed as an effort to approach communication in all its complex and heterogeneous facets, and to study a social interaction as it occurs, a kind of "natural history", a forceful data collection and exploration step that precedes theory building (Bateson, G., 1971). The team conducted a painstakingly detailed coding of multiple and simultaneous streams of interactive behavior, specifically overt communication behaviors: language, para-language, and body motion (kinesics). This project was the first major multimodal microanalytic study, contributing to new descriptive systems for para-language (Tragger, 1958) and body motion (Birdwhistell, 1952), analogous to structural linguistic systems.

Influences on the project mirror much of the intellectually interdisciplinary environment of the Macy Conferences, and ranged from Freudian theory, Gestalt psychology, social behaviorism, interpersonal psychoanalysis, cultural relativism, ethology to cybernetics, information theory, general systems theory and cinematography. These influences helped shaping what makes microanalysis a distinctive method. We unpack the three main aspects of this microanalytic approach and respective influences: its (1) object of analysis – the intact social interaction; its (2) context of application – natural environment; and the (3) level of detail – microscopic (at or below the scale of one second).

Object of Analysis

One of the most relevant aspects of the microanalytic method is its focus on social interaction**.** It reflects a change in the scope of behavioral sciences, from individual behavior to social interaction, that could be traced back to social behaviorism. Mead (1934) argued that the self develops as a product of social interaction, an influential view in the interpersonal theory of psychiatry of Sullivan (1955), and the birth of interpersonal psychoanalysis. Much of the initial microanalytic research, led by psychiatrists, followed that influence, and centered on the role of social interaction in the therapeutic process, and in developmental psychopathology (Bull, 2002). Social interaction began to be conceptualized in light of the recent theoretical frameworks derived from cybernetics, information theory, and general systems theory, and how they could be applied to human communication (Ruesch & Bateson, 1951), a view that the cognitive revolution would later neglect, with its focus in the information processing processes inside the individual (Gardner, 1987).

NHI researchers adopted an interpersonal deterministic approach that adapted some of the premises of Freudian theory to the ideas that interactional partners have limited conscious access to what happens throughout the communication process, and that everything that occurs is non-accidental and meaningful in the process of interaction. To decode this meaning, though, rather than focusing on the unconscious psychic forces in line with a mental deterministic approach, the microanalytic researcher should examine overt behavior and its perception. Adopting a Gestalt perspective, meaning can be derived from the analysis of the stream of communication, the events that punctuate it, and the contextual background to it. In light of social behaviorism and interpersonal psychoanalysis, instead of a synthesis of individual processes, microanalysis should aim at the interconnection between interpersonal processes and their context. The rules of interaction within a given interpersonal system would determine the transmission and reception of messages, as well as their distortions and the interpersonal problems that may arise. These communication failures have their pathogenesis in the continuous process of learning to communicate. But rather than ending in the chain of stimulus and responses as described by the, at the time, mainstream behaviorist model, learning is conceptualized in light of auto-regulated systems as described by cybernetics, relying on feedback for second order ("learning to receive signals") and third order ("learning to learn to receive signals") learning processes (Bateson, G.1971).

Context of Analysis

The objective of microanalysis is to analyze social interaction as it is. This emphasis is in opposition to previous approaches that privileged simplified models of how communication should be more efficient, or that analyzed communication from hypothetical or non-transcripted accurate samples (Weakland, 1967). Naturalistic observation of human communication was undoubtedly an objective of the NHI project. Microanalytic research borrowed much of the observation and recording techniques from fieldwork studies in ethnography and ethology (Bull, 2002).

The influence of Boas' ethnographic perspective on cultural anthropology and American structural linguistics accounts for the prominent prevalence of fieldwork in the development of both disciplines (Leeds-Hurwitz, 1987). The ethnographic work of Bateson and Margaret Mead, Boas' student, has one of the first examples of analysis through photographic material (G. Bateson & Mead, 1942). Another of his students, linguist Sapir, which have also influenced Sullivan's interpersonal thought, led the native American linguistic research. Privileging fieldwork and speech analysis over the study of written language, while utilizing tape recorded audio as the data for analysis. The use of direct observation and sound recording alerted American structural linguists to other aspects of communication, besides language, that should also be analyzed and later promoted the emergence of detailed analysis of other channels of communication (Bull, 2002; Leeds-Hurwitz, 1987; McQuown, 1971).

The emergent discipline of ethology also favored observation of animal behavior in their natural environment, pioneering some of the first techniques for stealth video recording interactional behavior (Bull, 2002). The focus of the ethological approach in description and classification became a model for the rigorous analysis of nonverbal behavior in microanalytic research. The work of ethologists Lorenz and Tinbergen also influenced Goffman – all of them participated in the Macy Conferences on group processes –, who referred to his fieldwork as human ethology (Bull, 2002). Akin to the ethnographic work of Chicago School on urban sociology, he used participant observation to analyze the mundane everyday face-to-face interactions (Goffman, 1967) in a common natural setting – public space (Goffman, 1963).

Level of Analysis

Microanalysis aimed at a fine-grained measurement of behavior, both spatially and temporally. The idea that any overt behavior – even its absence – could be communicative, motivated serious work in the analysis of different channels of communication. The detailed descriptions of speech and body motion benefited from the systematizations made by structural linguists, and their influence on the study of other communication modalities. While working with structural linguists Hocket and Smith, at the Foreign Service Institute of the Department of the State, Tragger, Birdwhistell, and Hall begin to develop systematic methods to analyze communication modalities that became a major part of the analytical work in the NHI: paralinguistics (Tragger, 1958); body motion – kinesics (Birdwhistell, 1952); and the use of space – proxemics (Hall, 1963).

These systematizations represented some of the initial efforts to understand how the stream of interactional behavior could be segmented and hierarchically organized from the most microscopic units of behavior to the more macroscopic events that punctuate the interaction. The fine-grained level of detail was made possible only by the evolution of recording technologies and cinematographic techniques. The use of slow motion, stop frame and rewind, enabled a level of control over the reproduction that was not possible before just through direct observation (Bateson, G., 1971; Bull, 2002; Leeds-Hurwitz, 1987). The final result was an extensive manual with guidelines for recording and manipulating audio and video interactional material, and descriptive systems for segmenting and coding verbal and nonverbal interactional behavior (McQuown, 1971).

Foundational as it was, the limitations of the NHI are evident. First, the sample was exceedingly small: a single interview with a family in their home environment. Second, coding was a time-consuming and resource demanding process; the multiple contributing teams took several years to produce a final report. Finally, it was above all a descriptive effort, and generalizing from the observed relations between different streams of behaviors was, at best, difficult. The obstacles in this approach remained in later successful applications of the microanalytic method; for example, ground-breaking studies in developmental psychopathology, that demonstrated the predictive power of interactive contingencies in mother-infant interactions, in the infant's future attachment style, took the team years to code (e.g., Beebe et al., 2010; Jaffe et al., 2001).

Microanalysis in Developmental Research

While the first publications of researchers involved in the NHI project exposed the methodological innovations to others (Condon & Ogston, 1967; Kendon, 1967; Scheflen, 1965), the evolution of recording and reproducing techniques made them more accessible to the developmental researcher (Beebe, 2014). To the early innovators, it became clear that mother-infant interaction unraveled at a split-second scale, across different channels of behavior, and that only by studying mother-infant interaction as it naturally occurs, one could access enough variation in order to understand the behavioral repertory of both interactional partners (Stern, 2002) – new tools would be required to describe the multimodal nature of mother-infant interaction, akin to those of naturalistic field research (Stern, 2002). Infancy researchers developed techniques of frame-by-frame analysis to code the microdynamics of different behavioral streams within mother-infant interaction on a second-by-second basis. The application of the microanalytic method also helped to unveil the ongoing process of adaptation between mothers and their infants. A key question to microanalytic developmental researchers became how to understand the temporal relation between interactional behaviors (Fogel, 1977), as well as the direction of influence between partners (Cohn & Tronick, 1988).

Early Microanalytic Studies of Mother-Infant Interactions

The first studies applying microanalytic methods to developmental research began in the late 1960s and early 1970s (Bateson, M. C., 1971; Brazelton et al., 1974; Condon & Sander, 1974; Stern, 1971). In 1971, the first two developmental microanalytic studies were published by M. C. Bateson (1971) and Stern (1971). Prior to this, some of NHI researchers published key studies (Condon & Ogston, 1967;

Kendon, 1967; Scheflen, 1965, 1967), and among them, Condon championed a major modification to the method that became the basis for Stern's and Condon's later developmental work (Condon & Sander, 1974; Stern, 1971, 1974). While the NHI microanalytic method focused on detailed reviews of extensive recorded material, Condon and Ogsten (Condon & Ogston, 1966, 1967) proposed the detailed frame-by-frame analysis of even shorter periods of interaction. Similarly, Brazelton, Koslowski and Main (Brazelton et al., 1974), used a variation on the method by describing longitudinal samples of 1 minute.

Stern's own retrospective views of this period are quite revealing: there is a place for description that precedes and informs experimentation; (largely inspired by Goffman), studying naturalistic interactions was vital to grasping the large range of mother and infant interactional behaviors and understand their dynamics; finally, the metaphor of a *dance* between mothers and their infants should be seen in the context of collaborations with dancers and choreographers that developed reproduction techniques to analyze body movement (Stern, 2002). Using this new window to mother-infant interactional behavior, Stern and colleagues revealed that most of it unravels at a split-second scale, through a multimodal conversation of, mostly, nonverbal behaviors, comparable to those studied in ethology.

While most of the developmental research of the 1970's organized around competencies strictly of the mother or the infant, microanalytic studies began to tackle with the structural and temporal organization of mother-infant interaction, documenting the interpersonal context where those competencies were actually used, and how they might be mutually regulated by both partners (Pérez & Español, 2016; Stern, 2002).

Next, we review microanalytic research for this early period, using a four-part structure that reflects the main questions of interest: (1) channel, (2) form, (3) structure, and (4) time.

Multimodal Behavioral Channels

Mother-infant interactions unravel through different channels of communication. Initial developmental microanalytic studies departed from other adult-infant face-to-face interaction research, by extending the channels analyzed (Pérez & Español, 2016).

M. C. Bateson (1971, 1975) introduced the study of proto-conversations by analyzing exchanges in vocalizations between mothers and their infants. Stern (1971, 1974) started to tackle the relations between eye contact and head and body orientation. While Condon and Sander (1974) studied the relation between mother's vocalizations and infant's type of movement. Later, Beebe and Gerstman

(1984) found evidence that multimodal parental behaviors work as a package of stimulation for the infant's level of involvement.

While other vocal and kinetic variables were, progressively, considered within developmental microanalytic studies, by the end of the 1980's, in part due to technological advances in speech analysis and a stagnation in kinetic technologies, vocalization became a predominant object of analysis, while the analysis of kinetic variables, even when contemplated on the studies, remained unpublished (Jaffe et al., 2001) – but see the work of Fogel and Thelen (1987) that sought to provide a theoretical framework, dynamic systems theory, to close the gap between the developmental study of movement, and the study of mother-infant interaction.

In the last decades, an increasing number of studies that walk the line between these two research traditions, began bringing the innovations of the 21st century to the studies of caregiver-infant interaction (e.g., de Barbaro, Chiba, & Deák, 2011; Smith, Yu, & Pereira, 2011).

Another channel diversification in the analysis of the stream of interactional behavior was the study of touch in the context of mother-infant interaction. Although, the relevance of tactile stimulation had been established, only by late 80's, early 90's research on mother-infant interaction through touch became more prevalent (Stack, 2004). Some of the first studies used adaptations of Tronick's (1978) still-face paradigm to understand the contributions of touch to mother-infant interaction (Gusella et al., 1988). This line of research has shown that touch can moderate the distressful effect of manipulations on voice and facial expression (Peláez-Nogueras et al., 1996; Stack & Muir, 1990, 1992); that infants are sensitive to touch manipulations (LePage, 1998; Stack & LePage, 1996) and explored the effects of touch and gesture integration (Arnold, 2002). Tronick's (1995) insights into the meaning of different touch patterns has also stimulated the construction of behavioral scales for coding touch patterns (Stack et al., 1996; Stepakoff, 1999).

Finally, an important addition to multimodal mother-infant interaction research was the ability to go beyond overt behavior through the introduction of measures that are correlates of psychological processes: physiological variables (Feldman, 2003; Feldman et al., 2011); neuroimaging measures, in particular the hyperscanning technique using dual electroencephalography (EEG) or functional near-infrared stereoscopy (fNIRS) (Nguyen, Bánki, et al., 2020; Wass et al., 2020). The combination of behavioral data with neural and physiological data also added a faster time scale, extending the behavioral time scale of the microanalytic approach; see Hoehl and Bertenthal (2021) for a recent discussion of the interrelations, and in some cases of dissociation, between behavioral and neural activity in mother-infant interactions.

We have described so far how microanalytic research has evolved to consider different channels of behavior. But while the ability to measure and code different streams of behavior may have evolved through time, this has more often resulted in specialization than actual integration, in contrast with the original proposals for microanalysis as inherently multimodal and interactional (G. Bateson, 1971; McQuown, 1971). Nevertheless, research constraints may ask for a deliberate decision regarding the behavioral units of analysis, and their integration, at the individual or dyadic level. Some of the best examples of an integrative effort in classical microanalytical research are revealed in the design of coding systems that enable the combination of individual behaviors, such as the monadic phases coding system (Tronick et al., 1980), or the coding of dyadic states using combinations of a specific behavior, such as Stern's approach to coding gaze (Stern, 1974). Still, new solutions for multimodal analysis of different streams of behavior at the individual and dyadic level are needed, that consider the greater level of detail that new recording methods have brought to microanalysis and provide complementary methods for behavioral pattern detection. Jaffe et al. (2001) offers an excellent example of how dyadic states can automatically be derived from the computation of individual vocalizations, using the Automatic Vocal Transition Analyzer (AVTA) model (Cassotta et al., 1964), providing integrated continuous and categorical data on dyadic turn-taking.

Form: The Abstract Shape of Interactive Behavior

Variation in the form of different interactional behaviors, across modalities, also became a focus of microanalytic research. One of the first insights of Stern's observations was regarding the form (shape) of maternal behaviors (Stern, 2002). The mothers' repertoire of interactive behavior, particularly vocalizations, had pitch, intensity, melody, and tempo characteristics that were qualitatively different to those expected in adult interactions; exaggeration and repetition appeared to be mechanisms to stimulate and engage infants in the interaction. Stern proposed the concept of temporal forms to encompass such variations and suggested they could be transversal to the quality of other channels of interactive behavior. Microanalytic research has brought attention to this phenomenon of caregiver's speech, today known as motherese, and its kinetic counterpart, motionese.

In 1975, Tronick and colleagues developed the still-face paradigm, and experimentally demonstrated the disturbing effects for infants of presenting no variation in mother's behavioral response in face-to-face interaction (Brazelton et al., 1975; Tronick et al., 1975). While Stern and collaborators focused only on the prosodic aspects of maternal speech (e.g., Stern, Spieker, Barnett, & Mackain, 1983; Stern, Spieker, & MacKain, 1982), Malloch, Sharp, Campbell, Campbell, and Trevarthen (1997) offered a variety of acoustic measures to study the dyadic patterns that emerge from prosodic variation within mother-infant interaction.

The Hierarchical Organization of an Interaction

Pauses and repetitions became the basis to distinguish between events that structure the interaction. Microanalytic research brought light to this structure and how it is hierarchically organized at different time-scales. In Stern (1974), maternal acts were proposed as the most elemental unit of mother-infant interaction, that could be further organized in a hierarchy of playful sequences, and full play episodes. Similarly, Stern, Beebe, Jaffe and Bennett (1977), suggested an equivalent to verbal communication – the phrase – as the first organization structure for grouping either vocalic and kinetic interactional behaviors, and showed how those phrases were progressively organized in larger units: sequences as repetitions of phrases, and episodes as clusters of sequences.

Fogel (1977) analyzed the temporal distribution of behaviors in mother-infant interaction, by distinguishing between runs and time-outs, as a binary on-off logic that temporally organizes the structure of mother-infant interaction. Runs are sequences of similar events and time-outs the intervals between them. This enabled the distinction of different levels of temporal organization considering differences in time-scale of each event.

The Temporal Relation Between Behavioral Streams

The temporal relation between caregiver and infant's behaviors is a core dimension of any interaction. Introducing the concept of proto-conversations, for describing the vocal exchanges between mothers and their infants, M. C. Bateson's (1971, 1975) work emphasized the relevance of studying infant pre-verbal vocal behavior on the context of mother-infant interaction, suggesting that a pattern of alternation organized the temporal structure of vocal exchanges since the second month of life.

By 1973, Joseph Jaffe's team began to apply models of dyadic coupling that analyze the temporal organization of adult verbal conversations to the analysis of mother-infant gaze behavior, suggesting similarities in the turn-taking organization of both *conversations* (Jaffe et al., 1973). Brazelton et al. (1974) used the concept of synchrony to refer to the temporal relation between mothers' behavior and infant's needs for attention. Condon and Sander (1974) provided evidence of synchrony between changes in mother's vocalizations and the movement (type, joints) of their infants, although this study's findings later became controversial (Gatewood & Rosenwein, 1981; McDowall, 1978a, 1978b).

Stern, Jaffe, Beebe and Bennett (1975) also found evidence of both simultaneity and alternation in infant's vocalization, as well as synchrony between smile and eye contact exchanges, leading to the conclusion that both organizations could be found in different interactive behaviors. In their 1977 study, both vocal and kinetic phrases exhibit non-exclusive patterns of synchronicity (Stern et al., 1977). Fogel (1977) demonstrated different types of temporal organization between different modalities of behaviors, suggesting that multiple temporal organizations could be found in mother-infant interaction and that the discussion of which would be more prevalent could be rather fruitless. In line with this, Beebe, Stern and Jaffe (1979) reinforced the importance of both simultaneous and alternation patterns of temporal organization.

Advantages and Limitations of a Microanalytic Approach to Developmental Research

The microanalytic approach to developmental research brought fundamental changes to how we study mother-infant interactions. First, it was a definitive departure from the retrospective conception of infant development offered by psychoanalysis, by bringing the study of infant's social and emotional development to the developing infant and the interpersonal context were those processes emerge. Second, microanalysis enabled the researcher to see mother-infant interaction for what it is. The value of an observational and descriptive approach that precedes theorization or experimentation is the ability to let the data inform us of the phenomenon as it is, before trying to explain it or control it. In doing so, microanalysis was able to start unpacking the microscopic units and dynamics, that were embedded in the more macroscopic infant or mother competencies that were starting to be explored experimentally, for example in studies of mother sensitivity or the effects of parenting quality (Ainsworth & Bell, 1974). Most microanalytic studies, directly or indirectly, speak to the question of adaptation between infants and their caregivers by starting from the microscopic analysis of what really happens in intact interactions. The majority of research presented earlier addressed aspects of stimulation, engagement, and regulation, aspects of caregiver and infant behavior that are indicative that both partners are adapting to each other. Although some studies focused on the influence of maternal behavior (Beebe & Gerstman, 1984), and others focused on the influence of the infant in generating those behaviors (Brazelton et al., 1974), the work of Cohn and Tronick (1988) empirically reinforced the bi-directionality of influence in mother-infant interaction. In fact, one of the most relevant outcomes of microanalytic research was to show that patterns of mother-infant interpersonal adaptation, as early as 3 to 4-months-old, can predict attachment style and cognitive development at later ages (Beebe et al., 2010; Beebe & Steele, 2016; Feldman & Greenbaum, 1997; Feldman, Greenbaum, Yirmiya, & Mayes, 1996; Jaffe et al., 2001). Unpacking these behaviors and dynamics has shown that microanalysis can help us not only understand the development of social and emotional processes in the infant, but also detect maladaptive patterns and intervene in early relationships, long before they can even be experimentally assessed.

Nevertheless, despite the advances of a microanalytic approach to the developmental study of early caregiver-infant interactions, some of the early limitations of the method remained while new issues appeared. First, although there was an initial preference for naturalistic interactions, most studies were conducted in laboratory environments that had some of the properties of natural environments; laboratory tasks provide valid correlates of developmental outcomes (e.g., Tamis-LeMonda et al., 2017) but are significantly different environments from where mother-infant interactions occurs; see discussion in Rogoff (2003) and Rogoff, Dahl, and Callanan (2018). Second, sample size was a limitation of most of the first microanalytic developmental studies, with a focus on case-studies or small samples. Third, with later increases in sample size, the variety of channels explored in the microanalytic studies decreased, (e.g., Jaffe et al., 2001) – but see recent work in Beebe et al. (2010). While never as integrative as the original project, early microanalytic developmental research often analyzed different modalities of behavior, while later research began to focus on specific channels of behavior. Fourth, detailed frame-by-frame coding remained a time-consuming and resource demanding process, which may explain how increases in sample size made coding of multiple behavioral streams infeasible, and vice-versa. An interesting trend of classic developmental microanalytic research, which could explain the progressive focus on vocalization research, was the use of automation to facilitate coding process, (e.g., Jaffe et al., 2001). Finally, although there were successful attempts to analyze the temporal structure and direction of influence within the interaction, there are limitations to using global measures of statistical association and additional requirements for the application of linear modelling techniques in time series analysis. Microanalytic studies in developmental research, after this early period, were also less frequent compared with other approaches. We argue that this derives from more than just the barrier imposed by the high cost of a single study, as there was also a broad turn towards the computational theory of mind and information processing that placed less emphasis in the relations between individual, social partners, and socio-cultural context (Gardner, 1987; Hutchins, 2010).

Evolution of Microanalytic Techniques

Interactive behavior spans across multiple modalities. While the first developmental microanalytic studies often privileged case-studies, and focused on a single behavioral modality, increases in sample size and in the modalities examined, imposed strenuous demands on the researcher using classic observational and coding techniques. Here we review some of the innovative techniques that are currently available, specifically: (1) sensing instruments and techniques to measure and process motion and physiological activity; (2) exploratory techniques to detect patterns within streams of high-density multidimensional data; and (3) statistical methods to model and test the contingencies between time series.

Sensing Techniques

Original microanalytic techniques relied on the available technologies, mainly audiovisual, to capture and reproduce the recorded behavior. Further developments on microanalytical developmental research, around the 80's and 90's, saw a progressive focus on vocal behavior (Pérez & Español, 2016), an effect of the ease of automating this modality (Jaffe et al., 2001).

Today, not only digital technology has progressed to allow capture, storage, and reproduction of increasing amounts of data, but the tools available for motion tracking and physiological recording have evolved considerably (Cornejo et al., 2017); there are now non-obtrusive ways to collect continuous data, with a high spatial and temporal level of detail, and automate some of the most arduous segmentation and coding processes.

Motion Tracking

Motion sensors are now an ubiquitous low-cost technology, since Inertial Measurement Units (IMUs) are incorporated in devices that we use in our everyday life. IMUs typically contain accelerometers and gyroscopes. Accelerometers can measure a body's acceleration up to three dimensions, and in some systems speed and position can be computed; gyroscopes can provide relevant information about a body's orientation. Because of their small form factor and cost, multiple sensors can easily be attached to the bodies of both interactional partners, without constraining their movement (Cornejo et al., 2017).

Motion capture systems can also use magnetic or optical technology to detect motion (these technologies were prevalent before low cost inertial sensors). Magnetic motion capture relies on the application of several magnetic sensors to a body, usually in a form of a suit; optical motion capture uses infrared cameras to detect markers attached to the body (Cornejo et al., 2017). Both require a level of technological apparatus that may remove participants from the conditions of a naturalistic interaction – e.g., suits, markers. Marker-less motion capture, though, based on multiple video recording and computer vision algorithms to detect trajectories in skeleton reproductions, may be a less intrusive alternative to capture motion from naturalistic interactions (Joo et al., 2015).

Computer vision has provided attractive solutions to the problem of measuring body motion without excessive instrumentation on the participant. A simple and highly efficient strategy is motion energy analysis (MEA), which relies on the difference in pixel color between sequences of frames to calculate movement dynamics, thus providing a measure of the degree of global change in movement over time; and when applied to specific regions of interest, it is possible to discriminate movement in a particular part of a body (Ramseyer, 2020); see Tschacher and Haken (2019) or Tschacher, Rees, and Ramseyer (2014) for a review of its use in adult clinical psychology studies. A more recent tool is OpenPose, a system designed for body pose estimation from unconstrained videos. Its use in infancy is recent but promising since it opens the possibility of analyzing a large range of video recordings, not just highly controlled environments (e.g., Long, Kachergis, Agrawak, & Frank, 2020; Sakurada et al., 2019).

Head-Mounted Cameras and Eye-Trackers

While a substantial amount of experimental work in developmental science is derived by examining patterns of infant gaze, studying the microdynamics of gaze in caregiver-infant naturalistic interactions requires an adaptation to standard eye-tracking devices which in turn requires infants to sit still in front of a monitor. Head-mounted eye-trackers can track gaze directly from the head of the participant allowing measurement of infant and caregiver eye gaze while they interact, without compromising movement, e.g. (Franchak et al., 2010; Yu et al., 2019; Yu & Smith, 2016). One caveat is that the technology has advanced considerably, but in infant studies the precision is still not equivalent to an external eye tracker, specifically when measuring the dynamics of fixations, see Hoehl and Bertenthal (2021) for an extensive discussion.

Another alternative is head-mounted cameras (Pereira et al., 2014; Yoshida & Smith, 2008); these capture the available visual scenes from the infant's perspective. This technique is a low cost solution to tap into the statistics of the infant's visual experiences and the infant's cognitive and social ecology; see the most recent findings reviewed in (Smith et al., 2015, 2018).

Physiological Recording

Measures of autonomic nervous system activity, such as heart rate or electrodermal activity, can provide alternative measures of affect, cognition, and behavior (Andreassi, 2007; Boucsein, 2012; Cacioppo & Tassinary, 1990). These measures thus open an unobservable dimension to the study of the microdynamics of caregiver-infant interaction, particularly in what concerns the interpersonal coordination within dyads. Electrophysiological recording devices rely on measuring the electrical activity on body tissues, requiring the application of electrodes or sensors to the specific body parts that are being measured. In a standard apparatus, this restricts participants' movements. Despite this unlikely scenario for research on naturalistic interacts, with the incremental increase in portability of the recording systems, there have been successful microanalytic studies using electrophysiological measures (Creavy et al., 2020; Feldman et al., 2011). Also, developments in wearable technology and the continuous drop in cost of sensors have produced new alternatives such as wristbands with electrophysiology recording (e.g. cardiac and dermal activity in the Empatica E4 product; Ragot et al., 2018), or a recent open-source system that consist of a wireless sensorized vest that infants can wear with cardiac, respiratory, and movement sensors (Maitha et al., 2020).

Hyperscanning

In the last two decades a new type of studies focused on inter-brain activity (brain-to-brain coupling) in social interaction has emerged as hyperscanning techniques became available (Czeszumski et al., 2020). Hyperscanning research uses the same neuroimaging or electrophysiological techniques used in individuals, such as electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and more recently, functional near-infrared stereoscopy (fNIRS) (Czeszumski et al., 2020). Not all of these instruments are adequate to the study of naturalistic interactions (MEG and fMRI problems are evident); nevertheless, EEG and fNIRS devices have progressed to a more portable form factor, and researchers have developed creative experimental designs and new data analysis techniques that can deal with the main problem, which is motion artefacts. Examples of empirical research and protocols are recently emerging in that direction (Leong et al., 2017; Markova et al., 2019; Nguyen, Bánki, et al., 2020; Nguyen, Schleihauf, et al., 2020; Reindl et al., 2018, 2019; Wass et al., 2020).

An interesting trend, that connects the development of motion tracking and physiological recording instruments, is a progressive shift to integrate biosensors in a wearable and wireless form factor to record and monitor physiological and motion activity in everyday life environments (Haynes & Yoshioka, 2007; Smith et al., 2018; Trull & Ebner-Priemer, 2013), these instruments are specifically tailored to overcome the restraints that standard recording and measuring instruments may impose and are already being successfully adapted to be used by infants in developmental research (de Barbaro, 2019; Maitha et al., 2020), see section five of this paper.

Exploration Techniques

As Bateson and Stern argued persuasively, when studying the intricate temporal structure of interactive behavior, hypothesis generation should precede theory building and hypothesis testing. Microanalytic studies generate datasets where a single dyadic interaction is captured by one or multiple video recordings, one or multiple audio recordings, and after data processing, a multivariate time series that includes categorical and/or quantitative variables. Exploratory techniques can uncover patterns and regularities from raw data and reveal something new about our phenomena of interest, which may in turn guide the generation of hypothesis, and the selection of adequate analytical tools for modeling and hypothesis testing.

Visual Analytics (Information Visualization)

Visualization of high-density raw data at an early stage of analysis can uncover patterns and regularities that increase the insight of researchers into the granularity of multimodal data and reveal underlying structure that is not bounded to theoretically prespecified measures (de Barbaro et al., 2013; Xu et al., 2020). The importance of visualization in exploratory data analysis was aptly summarized by Tukey: "*Pictures that emphasize what we already know [...] are frequently not worth the space they take. The greatest value of a picture is when it forces us to notice what we never expected to see.*" (Tukey, 1977, p. v).

Nevertheless, analysis of multimodal recordings of mother-infant interactions using visual analytics techniques are not common (Keim et al., 2010). One exception is Yu, Yurovsky and Xu (2012) work; they proposed a method of visual data mining, a human-in-the-loop process of pattern recognition, that rather than relying on automated analysis, leveraged on the perceptual abilities of human visual system, and the researcher's familiarity with phenomena of interest, in order to detect patterns in the multidimensional data set. The process follows a sequence of iterative steps: from (1) iterative visualization to (2) event-based exploration and (3) grouping and comparing of patterns.

Unfortunately, there are no generally accepted and available tools in the developmental community, tailored to the visualization of multivariate continuous and categorical data, although there are custom systems developed by researchers (e.g., Baur et al., 2020; Fouse, Weibel, Hutchins, & Hollan, 2011; Kim, Snodgrass, Pietrowicz, Karahalios, & Halle, 2013; Sousa et al., 2016).

A more automated solution to pattern recognition in multivariate data sets is the application of data mining algorithms. By converting multiple streams of continuous data to the bit level, information-theoretical measures can be applied to capture the information flow inside a network of behavioral streams (e.g., Choi, Yu, Smith, & Sporns, 2011).

Distributional Analysis

Analyzing the distribution of continuous streams of multivariate data sets can help us explore the temporal structure of such streams and guide the selection of subsequent analytical tools (Xu et al., 2020).

A useful distributional technique is burstiness analysis. Burstiness is a property in statistical physics that characterizes spike trains of activity, that can be quantified to analyze the distribution of individual interactional behaviors within and across modalities (Abney et al., 2018).

As described by Abney et al. (2018), the first step is for each individual behavior to create a binary spike train that represents the onset (and absence of an onset) of an event of interest; next, and for each modality, spike trains of individual behaviors are superimposed into a multivariate multimodal spike train. Inter-event interval distributions are computed from this multimodal series of events, by finding the temporal difference between two consecutive events. A burstiness parameter is estimated representing the combinations of bursts and lulls of a behavior.

Although with limitations to hypothesis testing, values of the burstiness parameter can be classified to distinguish between periodic, homogenous, and bursty processes (Abney et al., 2018; Xu et al., 2020).

Modelling Techniques

Since the first microanalytic studies, multiple techniques were used to model the temporal relations between behavioral modalities. The introduction of time series models, for example the influential method of Box and Jenkins (1970) to estimate the Auto Regressive Integrated Moving Average (ARIMA) class of regression models, was immediately put to use by the early innovators; e.g. Cohn and Tronick (1988), which shows how infants are also coordinated with the mother’s behavior in a form that is beyond simple frequency matching. Sequential analysis was also developed as a set of methodological techniques to analyze structure and temporality in the type of coded observational data encountered in microanalytic studies (Bakeman & Brown, 1977; Bakeman & Gottman, 1997; Jaffe et al., 1973)

From a statistics perspective, a powerful abstraction for interactive contingency is to model the multiple data streams of a dyadic interaction as a multivariate time series. We wish to measure (in a statistical inference sense) the association between any two pairs of the multivariate set across a range of temporal lags. Also, of importance to understanding interpersonal adaptation, we wish to measure the magnitude of interactive contingency for a specific dyadic direction; here, a predictive framework has been used frequently: if A and B are two social partners, and A adapts to B, the future behavior of A can be predicted by the past behavior B, above and beyond the information contained in the self-contingency of A – this is the definition used in Granger causality (Granger, 1969; Kirchgässner & Wolters, 2013).

Two components of interpersonal adaptation, however, add considerable complexity to modelling: (1) the multivariate time series includes both quantitative and categorical time series (e.g. body motion data vs. behavioral data – when coded within a stream as onset/offset events, behaviors are categorical time series); (2) the interaction is non-stationary (i.e. interactional data typically violates key assumptions of many stochastic processes).

To our knowledge, in the general case, these two challenges remain unsolved. Researchers have made great advances by carefully adding extra constraints and pursuing alternative approaches. For the multivariate quantitative/categorical problem, researchers have for example only considered the bivariate case, sometimes using more than one pair in the analysis; restricted the modelling to only the quantitative or the categorical domains; transformed a categorical time series into a quantitative one. Another recent alternative is to use information-theoretic measures such as transfer entropy (Bossomaier et al., 2016). For the non-stationary problem, researchers have explored several null-hypothesis testing techniques that are non-parametric; see Moulder, Boker, Ramseyer, and Tschacher (2018) for a review. Note, however, that social exchanges, including mother-infant interactions, tend to be locally stationary; for example, in the Cohn and Tronick (1988) study that demonstrated bidirectional coordination, the 3-minute interactions could be modelled with a stationary stochastic process.

In the following parts of this section, we describe a few methods that can assist in the issues we mentioned.

Correlational Analysis

Perhaps the most basic task, when measuring interpersonal adaptation, is detecting and quantifying a statistical association. In microanalysis, we wish to detect if in one, or several temporal lags, the behavioral stream of one partner is associated with the other partner, for instance: does an increase in the infant's object attention accompany a decrease in the mother's kinesic activity? Does an increase in the mother's level of engagement also lead to an increase in the infant's engagement? This amounts to computing some type of correlation but at different temporal distances. In the time-domain, for quantitative time series, a standard approach is the cross-correlation function (CCF): compute the Pearson correlation for all possible temporal lags inside a predefined window (e.g. indexing time by *t*, the correlation between the infant's head movement at time *t*, with the mother's head movement at time t, backwards in time *t-1*, *t-2*, etc, and forwards in time, *t+1*, *t+2*, etc,). Although the cross-correlation function has been widely used to analyze time series data from interactional partners (Delaherche et al., 2012), the CCF easily produces spurious correlations (Boker et al., 2002; Dean & Dunsmuir, 2016): if two stochastic processes are both auto-correlated, even if they are independent, the CCF can flag significant correlations where none exists; cf. an extensive discussion in Dean and Dunsmuir (2016). In behavioral data, the existence of autocorrelation is trivially present at a sufficiently short time scale. Moreover, while conventional parametric statistics requires data to be independent and identically distributed, and usually with constant parameters, interaction data rarely complies with those requirements; as pointed by Boker et al. (2002), behavioral time series rarely exhibit constant statistical properties over time.

We present two strategies: (1) removing auto-correlation; and (2) assuming local stationarity, compute multiple windowed cross-correlation functions. A third approach is to use non-linear techniques; we present Recurrence Quantification Analysis below as one example.

The first solution is referred to as pre-whitening. If a bivariate time series under analysis can be made to conform to the requirements of weak stationarity (e.g. usually by differencing), the auto-correlation can be removed by fitting an autoregressive model, to at least one of the time series, derive the residuals, and compute the CCF using the residuals instead. The resulting cross-correlation function can be safely interpreted (Dean & Dunsmuir, 2016). This is not without its disadvantages, since it typically imposes constraints when the variability of interactional behavior is the variable of interest (Boker et al., 2002).

The second strategy follows Boker et al. (2002) proposal of local stationarity, i.e. the stationarity in short durations of non-stationary processes. They proposed several extensions to the method of windowed cross-correlation in order to quantify variability of association, instead of removing the non-stationarity.

Windowed cross-correlation (Boker et al., 2002) breaks two time series in short intervals and cross-correlates these windows at differing times by lagging one window in relation to another, providing a moving estimate of the strength and lag of association between them throughout their course. In summary, instead of computing a single CCF using the entire time series, the time series is split in multiple windows (these can be allowed to overlap), a CCF is computed inside each window, quantitative properties of all CCFs are averaged to provide a single measure. Four parameters are required from the researcher to execute the analysis – window size, window increment, maximum lag, and lag increment – that are dependent on the phenomena of interest, and their choice should be theoretically informed and driven by exploratory analysis. Window size determines the duration of each interval as well as the number of observations in each window, and it is a critical choice for the assumption of local stationarity; it cannot be too small that it does not provide enough observations to estimate association, but should be small enough to capture short variations that reflect change and adaptation. Maximum lag determines the maximum interval between observations for each estimated association; it should be large enough to account for the expected delay in response between interaction partners, but not too long as to associate unrelated observations. Window increment determines the size of the shifts in observations from one window to another, as lag increment determines the size of shifts from one lag to the next. Larger values in both parameters may increase resolution at the cost of computational time (Boker et al., 2002; Moulder et al., 2018).

Global measures of association between time series can still be computed from windowed cross-correlations. In Moulder et al. (2018), measures derived from the pick peaking algorithm, a procedure suggested in Boker et al. (2002), and Fisher’s z transform, are both utilized to compare different methods for surrogate testing the significance of windowed cross-correlation. Notwithstanding, Dean and Dunsmir (2016) alert to the same dangers of spurious correlations when using windowed cross-correlations – comprised of many auto-correlated cross-correlations in sliding windows that are also by their sequential nature auto-correlated –, and to the necessity of supplemental procedures for testing the significance of such measures (see surrogate analysis, below).

Surrogate Analysis

As discussed above, spurious association is a problem when detecting dependency relationships between time-series (Dean & Dunsmuir, 2016). Surrogate analysis is an alternative to standard null-hypothesis testing to distinguish spurious from genuine dependency relationships, similarly to other bootstrapping methods such as randomization and permutation tests (Bernieri et al., 1988; Moulder et al., 2018; Reidsma et al., 2010).

Surrogate analysis uses generated surrogate data as a baseline for testing the significance of a chosen measure (Bernieri et al., 1988). The rationale behind surrogate testing is to identify a quality of interest, use a method to modify the data set in order to remove that quality of interest (without destroying other qualities of interest), and create a data set where the null hypothesis is certain to be true, then the same measure is applied both to the original data set, and to the surrogate data set to compare the likelihood of the real data statistic given a true null hypothesis (Moulder et al., 2018).

Several techniques for surrogate data generation can be employed, it is critical that the chosen method actually destroys the quality of interest and that its removal ensures that the null hypothesis is true. In Moulder et al. (2018) four methods are compared to destroy specific qualities of the original data related with temporal association: data shuffling, segment shuffling, data sliding, and participant shuffling. Beginning with the most destructive method, they can be applied in progression to successively attempt to reject more specific null hypotheses.

The most destructive method, data shuffling, shuffles all data points from each time series in order to create a new time series where none of the original data points remains in the same place. Albeit maintaining the same statistical distribution of the original, time dependency is destroyed – implying that the original data set will be tested against a null hypothesis of no time dependency between the original time series. A good measure of temporal association must pass this surrogate test to claim any time dependency between real time series, but given the physical constrictions on human behavior, shuffling continuous behavioral data also creates biologically impossible behavioral data where significance may be more informative of the differences between biologically constrained and biologically impossible systems (Moulder et al., 2018).

A less destructive method, segment shuffling, entails a segmentation of the original time series in short sections of the same size that are then shuffled until none of the original segments remains in the same place to generate the surrogate time series. Similar to the choice of window size in windowed cross-correlation, the segment size should be defined a priori by theoretical knowledge and exploratory analysis of the phenomenon – with implications to the ability of the significance test to reject the null hypothesis. Compared to data shuffling, the theoretically informed segments should retain some of the same properties of behavioral data, generating less biologically impossible data. It is also not a test against the absence of time dependency. Although some periodical behaviors larger than the segments may be affected by segment shuffling, data points within segments still conserve time dependency. Significance testing with this method tests instead against the null hypothesis of no time dependency between the segments in the original time series (Moulder et al., 2018).

Another method, data sliding implies only one cut on each of the original time series – preferably around the middle of the distribution – and the appending of the second part to the beginning of the first. With only one cut in the original data series it retains a lot of the same temporal structure with only one data point of biologically impossible behavior. While more vulnerable to spurious associations, this method is particularly informative to test for longer temporal dependencies and periodic behaviors against a null hypothesis of no influence of longer lags in time dependency between the original time series. While segment shuffling may be more appropriate to use with behaviors that occur in short bursts, data sliding may be better suited to test for global measures of time dependency (Moulder et al., 2018).

Finally, participant shuffling, requires the pairing of each element of a set of interactional partners with an element of another set of interactants, in order to create interactional surrogate data of participants who are not interacting with one another – pseudo-interactions. Significance testing is then made against the null hypothesis that the amount of time dependency in genuine interactions is no different to that expected on spurious interactions. A sufficient number of dyads (a reference number is 50) is required for the surrogate distribution to be adequate for significance testing. Given that no alteration is made in each of the original time-series to generate this surrogate data, it maintains the temporal structure of the original time series and is completely biologically plausible. Also, there could be similarities in the conditions of the genuine interactions. For that, rejecting the null hypothesis will be particularly harder than in the previous methods (Moulder et al., 2018).

Causality Analysis

A critical aspect of temporal dependencies between time series is to understand the directionality and magnitude of that dependency. Causality relations can be estimated from data, if we restrict causality to temporal contingency. Based on Weiner’s causality assumptions: the cause precedes the effect, and that there is unique information about the effect within that cause (Hlaváčková-Schindler et al., 2007), Granger (1969) operationalized a definition of causality based on the predictive value of past information for future occurrences. In this sense, a Granger causal relationship between variables A and B is determined when adding past information about variable A is more predictive of future events in variable B than the past information of variable B alone.

There are several methods to estimate granger causality for time series, both in time and frequency domains, that can be chosen according to their utility; see Cekic, Grandjean, and Renaud (2017) for a review, and Sun et al. (2004) for the application of the Wavelet Coherence Transform in neuroimaging, a frequency domain technique that is also being applied in studies of mother-infant social interaction (e.g., Nguyen et al., 2021). While Granger causality was first operationalized (for practicality) within a linear autoregressive framework, that assumed stationary time series, the original formulation did not exclude non-linear dynamics, and extensions to non-linear Granger causality have been proposed (Hlaváčková-Schindler et al., 2007). While matching the choice of model to the underlying dynamics used is critical to avoid spurious causality, a non-parametric alternative to non-linear granger causality, free from model mismatch, is to conceptualize causality in terms of information-theoretical measures such as transfer entropy – e.g. for a review of information-theoretical estimators see Hlaváčková-Schindler et al. (2007) and Bossomaier et al. (2016).

Recurrence Analysis

A final alternative is to use non-linear techniques for analyzing time dependency between time series data. A well-known technique is cross-recurrence analysis: it is the bivariate extension of Recurrence Quantification Analysis (RQA), and can be described as a non-linear generalization of the cross-correlation function. RQA is both a visualization tool, the recurrence plot – itself an exploratory data analysis tool (e.g., (Anderson et al., 2013) – and a set of quantitative measures on the recurrence plot, cross-recurrence quantification (Marwan et al., 2002; Marwan & Kurths, 2002; Shockley et al., 2002).

Cross-recurrence plots provide a graphical representation of the points in time where both systems show similar states or patterns of change – the recurrence points – and enables the identification of a set of structures informative of the dynamics between systems, such as stationarity and non-stationarity, cyclicity, fluctuations, determinism, chaos, embedding, and laminarity. With cross-recurrence quantification some of these dynamics identified in the cross-recurrent plot can be quantified (Marwan et al., 2007).

Cross-recurrence quantification analysis includes a set of measures that measure the dynamics in the patterns of recurrence. The simplest measure, *recurrence rate*, gives a global measure of the density of recurrences, and corresponds to the cross-correlation sum. Measures of the line structures in the cross-recurrence plot, though, can be used to quantify non-linear patterns that are undetectable by a cross-correlation function (Marwan et al., 2007).

From diagonal structures, only chaos-order transitions can be identified. *Determinism*, a measure of predictability can be calculated from the ratio of recurrence points that form the diagonal lines. The *average diagonal length* can be used as an indication of the mean prediction time. The *maximum length* of a diagonal line can be used to detect *divergence*. *Entropy* can be calculated to determine complexity. All previous measures, including *recurrence rate*, can also be calculated for each diagonal. These diagonal-wise measures can inform of the similarity between the dynamics of both systems over time. By applying measures of symmetry and asymmetry on the diagonal-wise measures we can also quantify the interrelations between the coupled systems and determine when one system leads another. Other measures that can be quantified from diagonal structures are *trend*, a measure of non-stationarity, and *ratio*, a measure of transitions in dynamics (Marwan et al., 2007).

From vertical structures, chaos-chaos transitions can also be revealed. The ratio of recurrence points that form horizontal lines can be calculate as a measure of *laminarity*. The average length of vertical structures is defined as *trapping time* and provides an estimate of how long a state will remain. The *maximum length* of the vertical lines can also be computed. For a thorough exposition of these and other recurrence analysis measures and their estimation, see (Marwan et al., 2007).

An extension of cross-recurrence analysis, for discrete time series, can be a good alternative to overcome the limitations of defining recurrences in terms of similar states. By operationalizing recurrence as behavioral matching (Cox et al., 2016), not only recurrences in similar states can be detected but also recurrences between potentially dissimilar states that the researcher chooses to match. Cox et al. (2016) propose two procedures to help with the tasks of visualizing different coded matches – chromatic cross-recurrence analysis – and to quantify the vertical and horizontal structures that are prominent with discrete data.

Contemporary Microanalysis

We selected two domains, recent users of the methodological innovations described earlier. Both examples, in particular the first one, egocentric vision – the study of the infant's visual ecology – are not framed in terms of microanalysis; we selected them for two reasons: first, all studies include or are concerned with social interaction contexts and aim at measuring the real-time dynamics of social exchanges with infant; second, they illustrate how changing the lens in the social microscope can provide new perspectives to caregiver-infant interaction, similar to the early period of microanalysis.

Egocentric Vision

Developmental studies of egocentric or first-person view (Franchak et al., 2010; Yoshida & Smith, 2008) began with the use of head-mounted cameras (Pereira et al., 2009; Yoshida & Smith, 2008; Yu et al., 2008) to capture the visual perspective of the infant and their caregiver in interaction, followed soon after by head-mounted eye trackers (Franchak et al., 2010). This new methodological approach had a similar impact to what recording and reproducing the interaction had to the early microanalytic researcher: the possibility of observing from a completely different angle and time scale (Smith et al., 2018). Studying interaction from a first-person perspective is in contrast with the standard third-person perspective, used in developmental research for decades, in the way that it captures the moment-to-moment dynamics of the infants' (and of their caregivers') visual experience as it relates to their own action (Smith et al., 2011).

Yu and Smith (2016) examined the role of social interaction in the development of sustained attention, a fundamental developmental process often believed to be a product of endogenous maturation of an infant's ability for attentional control. By using head-mounted eye-trackers to record the first-person gaze microdynamics in parent-infant free flowing play interactions, the authors discovered that one-year-old infants extend the duration of their visual attention to an object of play, when their parents attended to the same object (Yu & Smith, 2016). This suggests that parent-child joint attention may have a role in the development of sustained attention, challenging the endogenous perspective (Yu & Smith, 2016).

Developmental studies of egocentric or first-person view (Franchak et al., 2010; Yoshida & Smith, 2008) have become a recent paradigmatic illustration example of the impact of the social microscope; see Smith et al. (2018) for a recent review.

Daily Activity Sensing

The idea of a natural history of social behavior, a description of a phenomenon in the context where it actually occurs, was a major motivation for microanalysis. This ecological perspective, grounded in cybernetics and systems thinking, had also an impact on developmental research. Microanalytic studies of mother-infant interactions stressed the value of observing mother-infant interactions in their natural environment (M. C. Bateson, 1971; Stern, 1971).

However, most of developmental research, including microanalytic, has been made inside of the laboratory, with setting adaptations to resemble natural environments, often privileging scripted interactions (e.g. the structured social interaction task) to the richness of truly ecological interactions (de Barbaro, 2019). There is no lack of evidence, nonetheless, for the importance of considering the cultural and daily conditions of infants' natural environments in order to understand development (Rogoff, 2003; Rogoff et al., 2018; Tamis-LeMonda et al., 2017).

There are qualitative differences between the conditions of structured interactions, staged by the experimenter in a laboratory setting, and the conditions where interactions naturally unfold, in their environment, intertwined with everyday life routines. In a recent methodological study, Tamis-LeMonda et al. (2017) compared a structured interaction play task, to a naturalistic observation in home environments, in terms of infant's language production. The study found both differences and similarities in the results that could illuminate the benefits of each method. While structured interactions held similar results to peak language production periods in the home environment, the temporal dynamics of infant's language production were better captured in the naturalistic setting.

With the evolution of sensing instruments and techniques, in the direction of portability and wearability since the turn of the century (Abowd & Mynatt, 2000), collecting interactional data from naturalistic environments has become a less intrusive approach. From video and audio recording, to motion and physiological sensing, there has been enough innovation in ubiquitous computing to enable multidimensional data from everyday activity to be collected in unobtrusive and unscripted ways in the contexts where interactions naturally unravel (de Barbaro, 2019). When paired with adequate analytical and modeling techniques to understand this data it can throw new light on the complex dynamics of early life interactional behavior (de Barbaro, 2019).

In adult clinical psychology studies, ambulatory assessment (Haynes & Yoshioka, 2007; Trull & Ebner-Priemer, 2013) has already been making use of these innovations, collecting data from individuals in their natural environments. Innovations in developmental research are rapidly catching up by introduction of infant adapted technologies for daily activity sensing, as head-mounted cameras and eye-trackers (Franchak et al., 2010; Yoshida & Smith, 2008) or wearable integrated biosensing systems (Maitha et al., 2020).

In conclusion, daily activity sensing can provide access to behaviors that are not available in a laboratory setting, interactions that extend over time and even rare events (de Barbaro, 2019).

Discussion

We have considered the foundations of microanalytic research, and how transposing the method to the study of mother-infant interaction has brought relevant contributions to our understanding of infant's development, but also obstacles to the ability of the microanalytic researcher to go further. We reviewed how sensing, exploration, and modeling techniques have evolved to tackle those obstacles and how applying these techniques is already producing new findings. We will continue by discussing the tradeoffs of using the micro-analytical approach.

Advantages and Potentialities

Microanalysis is often compared to a social microscope, a tool to unveil what is not available to the researcher through observation of social interaction at its natural time-scale. The strength of microanalysis lies in its ability to help researchers go beyond their limitations, while still making the most of their strengths to identify patterns and relations. Both use cases we included here, egocentric vision and daily life sensing research, are illustrative of the potential that new sensing techniques have to capture new information about the structure of the infant's social exchanges with others. Even with the helpful evolution of the algorithms that automate some of the detection and classification of patterns, the transition from observing macroscopic behavioral patterns in interaction, to observing patterns in the high-density data collected is not straightforward. Using exploratory techniques for information visualization, such as visual data mining or the cross-recurrence plot, can assist in this effort. Additionally, visualization tools can be used to bring together the data and the observed interaction. Sometimes the questions we ask regarding interpersonal adaptation demand complex modelling techniques; we chose to review some of the techniques that better address the sequential and split-second nature of interactional events in order to derive relations between them. What those techniques provide are measures of temporal association between events that may be described as coordination, synchrony, or interactive contingencies – the moment-to-moment adjustments of a partner's behavior to the behavior of another, revealing the adaptation dynamics of the interaction (Burgoon, Stern & Dillman, 1995).

Although most of microanalytic research focused on the microscopic, frame-by-frame, split-second analysis of changes in the continuous stream of interactional behaviors, not all interactional phenomena occurs at the same time scale, even within co-occurring behaviors, depending on their channel of presentation and functional relation. We have also seen how attending to different time scales, can help us organize the interaction in hierarchical temporal structures, often punctuated by silences, repetition, and variation. This hierarchical structure reflects the intricate balance between stimulatory and regulatory longer processes, a foundation in the development of adaptive and maladaptive interpersonal patterns. Undoubtably, the tools we propose to enhance our social microscope can be used to capture data for longer durations, in the environments where longer scale interactional phenomena may occur, and to detect and model interactional phenomena across multiple time scales. On the flipside of this discussion are faster time scales than the split-second behavioral time scale. It is now possible to study naturalistic social interactions that include simultaneous hyperscanning (EEG or fNIRS) or physiological recording. The importance of this is to connect observable behavior with the underlying neural processes and start to disentangle moments where the observable behavior is similar but the neural signatures are distinct, indicative of different processes at work; a recent review of this link in the context of social attention is in Hoehl and Bertenthal (2021).

Finally, and in conclusion, the most urgent direction in microanalytic research is perhaps to probe into an essential property: the multidimensionality of the interactional phenomena, that is, understanding the relations between multiple streams of behavior, how they evolve over time, and outside of the laboratory. Two fundamental challenges within this direction are to continue improving our ability to automate data collection and to develop the appropriate tools to enable a meaningful integration of the different streams of behavior.

Limitations and Solutions

It is useful, for a pragmatic application of the microanalytic approach, to also consider its limitations and how they can or cannot be solved. The problems of ecological validity are inherently central to a microanalytic approach. Microanalysis emerged from the need to examine social interactions as they unravel in their natural environments. Still, from the beginning, the microanalytic approach also introduced something external to that environment, the technical apparatus for recording the interaction. With the advances in sensing techniques and their application to caregiver-infant interaction research, a new layer of technological apparatus is added to the context of data collection. To what extent these external additions prevent us from capturing the intact ecology of social interaction should also be considered. As far as the sensing techniques we emphasized and the devices that are being adapted to developmental research, the focus has been in mobility and wearability. A complementary direction is the development of laboratory spaces that better mimic the natural environment, while concealing some of technological apparatus.

Another issue in validity is the ability of automated solutions to provide reliable classifications compared to trained human coders. As far as we know, there is still little to no research that directly addresses this question; one exception is a recent study that compared automatic and manual measures of conversational turns with the Language Environment Analysis (LENA™) system for audio recordings (Ramírez et al., 2021). This study compared LENA system's algorithm for detecting turns – Conversational Turn Count (CTC) - with the turns detected by manual coders, and the findings point to the need for further research to validate the CTC metric (Ramírez et al., 2021). Thus more research is needed that tackles the validity of sensing systems when compared to manual coding. Nonetheless, we should bare in mind that most of the sensing techniques we have discussed provides data at a rate and level of detail that exceeds the possibilities of manual coding and can further complement the human-based coding, but require different approaches to data analysis.

A good illustration of both obstacles we have discussed is the applicability of haptic sensing techniques to the study of social touch in caregiver-infant interactions. While there is emergent research in mediated social touch, and in social touch in human-robot interactions using haptic technology – see Huisman (2017) for a review –, these haptic systems are still not adapted to use in naturalistic interactions. Motor behaviors that can be categorized as social touch are quite varied (Serra et al., 2020) and with current computer vision technology, only trained coders are able to extract them from a video recording.

Conclusion

Microanalysis stands as an example of a research tradition that has tackled the methodological challenges that interactional behavior poses to the developmental researcher. By focusing on the emergence of the microanalytic approach and the pioneering ideas beyond it, we intended to connect the context of microanalysis emergence with current developments, methodological and empirical.

The insights and questions that preoccupied the pioneers of microanalytic research are still relevant today to those researching the microdynamics of mother-infant interaction. As developmental science ventures outside the laboratory, we are reminded of the value of observing real life interactions as they unfold, at their natural rhythm and in their natural environments (Cychosz et al., 2020; de Barbaro, 2019; Rogoff, 2003; Rogoff et al., 2018; Tamis-LeMonda et al., 2017). The challenge of how to capture the complexity of multiple, simultaneous streams of behavior is still present, while some of the limitations of manual coding techniques were replaced by automated alternatives (Jaffe et al., 2001; Smith et al., 2018). Likewise, understanding the complex relations between multimodal streams of interactional behavior, requires tools to explore and visualize multivariate time series data, and appropriate techniques for statistical inference and modeling of interactional time series (Xu et al., 2020).

Another parallel we can trace between the past and present of microanalytic research is how changes in the lens used in the metaphorical social microscope have improved our understanding of mother-infant interaction (Franchak et al., 2010; Smith et al., 2018; Yu & Smith, 2016). In early microanalytic research, the introduction of video recording and reproduction techniques opened a window to explore the reality of mother-infant interaction; the introduction of new sensing techniques (Abowd & Mynatt, 2000) in a wearable and unobtrusive form factor, is already bringing new perspectives to our study of mother-infant interactions (de Barbaro, 2019). We presented two illustrations, taken from recent developmental research, egocentric vision, and daily activity sensing. The case of egocentric vision shows the potential of the integrated multimodal study of natural interactions. With daily activity sensing, the developmental researcher can gain access to multimodal data of mother-infant interactions as they unfold in their natural environment.

Although we now have the potential to overcome some of the most important limitations of previous microanalytic research, new challenges also arise from new solutions. First, it is not yet possible to completely automate the process of data collection and coding. Another relevant challenge is that of making sense of the high-density multidimensional data that can be collected using the new resources available (Xu et al., 2020). While there are computer algorithms that may help the search for patterns in the data, the balance between the benefits of data-driven discovery, and the challenges of segmenting the reality in theoretically and methodologically significant chunks, make exploratory tools with human-in-the-loop solutions still preferable (de Barbaro et al., 2013; Yu et al., 2012). Moreover, the variability over time within multivariate interactional time-series data challenges the statistical assumptions inherent to many of the procedures available for statistical inference and modelling, making the applicability of the most often used statistical solutions still not completely understood (Boker et al., 2002). Emerging non-linear techniques may be promising alternatives to the present challenges (Xu et al., 2020).

We argue that the microanalytic project is as relevant today as when it was proposed. Our main argument, of bringing new lenses to the social microscope, is not one of substitution. New methods have also introduced new challenges. We argue for the complementarity of different tools in the toolbox of the developmental researcher when approaching the complex dynamics of mother-infant interaction.

References

Abney, D. H., Dale, R., Louwerse, M. M., & Kello, C. T. (2018). The bursts and lulls of multimodal interaction: Temporal distributions of behavior reveal differences between verbal and non-verbal communication. *Cognitive Science*, *42*(4), 1297–1316. https://doi.org/10.1111/cogs.12612

Abowd, G. D., & Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction*, *7*(1), 29–58. https://doi.org/10.1145/344949.344988

Ainsworth, M. D. S., & Bell, S. M. (1974). Mother-Infant interaction and the development of competence. In K. J. Connolly & J. Bruner (Eds.), *The growth of competence* (pp. 97–118). Academic Press.

Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, *45*(3), 842–856. https://doi.org/10.3758/s13428-012-0299-5

Andreassi, J. L. (2007). *Psychophysiology: Human behavior and physiological response.* (5th ed.). Lawrence Erlbaum associates.

Arnold, S. L. (2002). *Maternal tactile-gestural stimulation and infants' nonverbal behaviors during early mother-infant face-to-face interactions: Contextual, age, and birth status effects.* [Doctoral dissertation, Concordia University]. https://spectrum.library.concordia.ca/1770/

Bakeman, R., & Brown, J. V. (1977). Behavioral dialogues: An approach to the assessment of mother-infant interaction. *Child Development*, *48*(1), 195. https://doi.org/10.2307/1128898

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction*. Cambridge University Press. https://doi.org/10.1017/CBO9780511527685

Bakeman, R., & Quera, V. (2011). *Sequential Analysis and Observational Methods for the Behavioral Sciences*. Cambridge University Press. https://doi.org/10.1017/CBO9781139017343

Bateson, G. (1971). Comunication. In N. A. McQuown (Ed.), *The natural history of an interview* (pp. 1–40). University of Chicago Library.

Bateson, G., & Mead, M. (1942). *Balinese character: A photographic analysis*. New York Academy of Sciences.

Bateson, M. C. (1971). The interpersonal context of infant vocalization. *Quarterly Progress Report of the Research Laboratory of Electronics*, *100*, 170–176.

Bateson, M. C. (1975). Mother-infant exchanges: the epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, *263*(1), 101–113. https://doi.org/10.1111/j.1749-

6632.1975.tb41575.x

Baur, T., Clausen, S., Heimerl, A., Lingenfelser, F., Lutz, W., & André, E. (2020). *NOVA: A Tool for explanatory multimodal behavior analysis and its application to psychotherapy* (pp. 577–588). https://doi.org/10.1007/978-3-030-37734-2_47

Beebe, B. (2014). My journey in infant research and psychoanalysis: Microanalysis, a social microscope. *Psychoanalytic Psychology*, *31*(1), 4–25. https://doi.org/10.1037/a0035575

Beebe, B., & Gerstman, L. (1984). A method of defining "packages" of maternal stiimulation and their functional significance for the infant with mother and stranger. *International Journal of Behavioral Development*, *7*. https://doi.org/10.1177%2F016502548400700302

Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., Bahrick, L., Andrews, H., & Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother–infant interaction. *Attachment & Human Development*, *12*(1–2), 3–141. https://doi.org/10.1080/14616730903338985

Beebe, B., & Lachmann, F. M. (2013). *The origins of attachment: Infant research and adult treatment*. Routledge. https://doi.org/10.4324/9781315858067

Beebe, B., & Steele, M. (2016). How does microanalysis of mother-infant communication inform maternal sensitivity and infant attachment? *Maternal Sensitivity: Mary Ainsworth's Enduring Influence on Attachment Theory, Research, and Clinical Applications*, *15*(5–6), 141–160. https://doi.org/10.1080/14616734.2013.841050

Beebe, B., Stern, D. N., & Jaffe, J. (1979). The kinesic rhythm of mother-infant interactions. In A. W. Siegman & S. Feldstein (Eds.), *Of speech and time: Temporal patterns in interpersonal contexts* (pp. 23–34). Lawrence Erlbaum.

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: measuring the entrainment process in mother-infant interactions. *Journal of Personality and Social Psychology*, *54*(2), 243–253. https://doi.org/10.1037/0022-3514.54.2.243

Birdwhistell, R. L. (1952). *Introduction to kinesics: An annotation system for analysis of body motion and gesture*.

Boker, S. M., Xu, M., Rotondo, J. L., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*, *7*(3), 338–355. https://doi.org/10.1037/1082-989X.7.3.338

Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016). An introduction to transfer entropy. In *An Introduction to Transfer Entropy*. https://doi.org/10.1007/978-3-319-43222-9

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). Springer Science, Business & Media.

Box, G., & Jenkins, G. (1970). *Time series analysis: Forecasting and control*. Holden-Day.

Brazelton, T. B., Koslwski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). Willey-Interscience.

Brazelton, T. B., Tronick, E. Z., Adamson, L., Als, H., & Wise, S. (1975). Early mother-infant reciprocity. *Ciba Foundation Symposium*, *33*, 137–154. https://doi.org/10.1002/9780470720158.ch9

Bull, P. (2002). *Communication under the microscope: The theory and practice of microanalysis*. Routledge.

Burgoon, J. K., Stern, L. A., & Dillman, L. (1995). *Interpersonal adaptation: Dyadic interaction partners*. Cambridge University Press.

Cacioppo, J. T., & Tassinary, L. G. (1990). *Principles of psychophysiology: Physical, social, and inferential elements*. Cambridge University Press.

Cappella, J. N. (1981). Mutual influence in expressive behavior: Adult-adult and infant-adult dyadic interaction. *Psychological Bulletin*, *89*(1), 101–132. https://doi.org/10.1037/0033-2909.89.1.101

Casillas, M., & Cristia, A. (2019). A step-by-step guide to collecting and analyzing long-format speech environment (LFSE) recordings. *Collabra: Psychology*, *5*(1), 1–21. https://doi.org/10.1525/collabra.209

Cassotta, L., Feldstein, S., & Jaffe, J. (1964). AVTA: A device for automatic vocal transaction analysis. *Journal of the Experimental Analysis of Behavior*, *7*(1), 99–104. https://doi.org/10.1901/jeab.1964.7-99

Cekic, S., Grandjean, D., & Renaud, O. (2018). Time, frequency, and time-varying Granger-causality measures in neuroscience. *Statistics in Medicine*, *37*(11), 1910–1931. https://doi.org/10.1002/sim.7621

Choi, H., Yu, C., Smith, L. B., & Sporns, O. (2011). From data streams to information flow: Information exchange in child-parent interaction experiment and data preprocessing. *Proceedings of the 33nd Annual Conference of the Cognitive Science Society*, 868–873.

Cohn, J. F., & Tronick, E. Z. (1988). Mother-Infant face-to-face interaction: influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology*, *24*(3), 386–392. https://doi.org/10.1037/0012-1649.24.3.386

Condon, W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior. *The*

Journal of Nervous and Mental Disease, *143*(4), 338–347. https://doi.org/10.1097/00005053-196610000-00005

Condon, W. S., & Ogston, W. D. (1967). A segmentation of behavior. *Journal of Psychiatric Research*, *5*(3), 221–235. https://doi.org/10.1016/0022-3956(67)90004-0

Condon, W. S., & Sander, L. W. (1974). Synchrony Demonstrated between movements of the neonate and adult speech. *Child Development*, *45*(2), 456. https://doi.org/10.2307/1127968

Cornejo, C., Cuadros, Z., Morales, R., & Paredes, J. (2017). Interpersonal coordination: Methods, achievements, and challenges. *Frontiers in Psychology*, *8*(SEP), 1–16. https://doi.org/10.3389/fpsyg.2017.01685

Cox, R. F. A., van der Steen, S., Guevara, M., Jonge-Hoekstra, L., & van Dijk, M. (2016). Chromatic and anisotropic cross-recurrence quantification analysis of interpersonal behavior. In V. Romero, P. Fitzpatrick, R. C. Schmidt, & M. J. Richardson (Eds.), *Recurrence Plots and Their Quantifications: Expanding Horizons* (pp. 209–225). https://doi.org/10.1007/978-3-319-29922-8

Creavy, K. L., Gatzke-Kopp, L. M., Zhang, X., Fishbein, D., & Kiser, L. J. (2020). When you go low, I go high: Negative coordination of physiological synchrony among parents and children. *Developmental Psychobiology*, *62*(3), 310–323. https://doi.org/10.1002/dev.21905

Cychosz, M., Romeo, R., Soderstrom, M., Scaff, C., Ganek, H., Cristia, A., Casillas, M., de Barbaro, K., Bang, J. Y., & Weisleder, A. (2020). Longform recordings of everyday life: Ethics for best practices. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01365-9

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., Schreiber, F., Rendon, Z. Z., & König, P. (2020). Hyperscanning: A valid method to study neural inter-brain underpinnings of social interaction. *Frontiers in Human Neuroscience*, *14*(February), 1–17. https://doi.org/10.3389/fnhum.2020.00039

de Barbaro, K. (2019). Automated sensing of daily activity: A new lens into development. *Developmental Psychobiology*, *61*(3), 444–464. https://doi.org/10.1002/dev.21831

de Barbaro, K., Chiba, A., & Deák, G. O. (2011). Micro-analysis of infant looking in a naturalistic social setting: Insights from biologically based models of attention. *Developmental Science*, *14*(5), 1150–1160. https://doi.org/10.1111/j.1467-7687.2011.01066.x

de Barbaro, K., Johnson, C. M., Forster, D., & Deák, G. O. (2013). Methodological considerations for investigating the microdynamics of social interaction development. *IEEE Transactions on Autonomous Mental Development*, *5*(3), 258–270. https://doi.org/10.1109/TAMD.2013.2276611

Dean, R. T., & Dunsmuir, W. T. M. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, *48*(2), 783–802. https://doi.org/10.3758/s13428-015-0611-2

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, *3*(3), 349–365. https://doi.org/10.1109/T-AFFC.2012.12

Feldman, R. (2003). Infant-mother and infant-father synchrony: The coregulation of positive arousal. *Infant Mental Health Journal*, *24*(1), 1–23. https://doi.org/10.1002/imhj.10041

Feldman, R., & Greenbaum, C. W. (1997). Affect regulation and synchrony in mother-infant play as precursors to the development of symbolic competence. *Infant Mental Health Journal*, *18*(1), 4–23. https://doi.org/10.1002/(SICI)1097-0355(199721)18:1<4::AID-IMHJ2>3.0.CO;2-R

Feldman, R., Greenbaum, C. W., Yirmiya, N., & Mayes, L. C. (1996). Relations between cyclicity and regulation in mother-infant interaction at 3 and 9 months and cognition at 2 years. *Journal of Applied Developmental Psychology*, *17*(3), 347–365. https://doi.org/10.1016/S0193-3973(96)90031-3

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4), 569–577. https://doi.org/10.1016/j.infbeh.2011.06.008

Fogel, A. (1977). Temporal organization in mother-infant face-to-face interaction. *Studies in Mother-Infant Interaction*, *January 1977*, 119–152.

Fogel, A., & Thelen, E. (1987). Development of early expressive and communicative action: reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, *23*(6), 747–761. https://doi.org/10.1037/0012-1649.23.6.747

Fouse, A. S., Weibel, N., Hutchins, E., & Hollan, J. D. (2011). ChronoViz: A system for supporting navigation of time-coded data. *Conference on Human Factors in Computing Systems - Proceedings*, 299–304. https://doi.org/10.1145/1979742.1979706

Franchak, J. M., Kretch, K. S., Soska, K. C., Babcock, J. S., & Adolph, K. E. (2010). Head-mounted eye-tracking of infants' natural interactions: A new method. *Eye Tracking Research and Applications Symposium (ETRA)*, *May 2014*, 21–28. https://doi.org/10.1145/1743666.1743671

Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution*. Basic Books.

Gatewood, J. B., & Rosenwein, R. (1981). Interactional synchrony: Genuine or spurious? A critique of recent research. *Journal of Nonverbal Behavior*, *6*(1), 12–29. https://doi.org/10.1007/BF00987933

Goffman, E. (1963). *Behavior in public places: Notes on the social organization of gatherings*. The Free Press.

Goffman, E. (1967). On face-work. In *Interaction Ritual* (pp. 5–45).

Gottman, J. M. (1990). *Sequential analysis: A guide for behavioral researchers*. Cambridge University Press.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791

Gusella, J. L., Muir, D., & Tronick, E. Z. (1988). The effect of manipulating maternal behavior during an interaction on three- and six- month-olds ' affect and attention. *Child Development*, *59*(4), 1111–1124.

Hall, E. T. (1963). A system for the notation of proxemic behavior. *American Anthropologist*, *65*(5), 1003–1026. https://doi.org/10.1525/aa.1963.65.5.02a00020

Haynes, S. N., & Yoshioka, D. T. (2007). Clinical assessment applications of ambulatory biosensors. *Psychological Assessment*, *19*(1), 44–57. https://doi.org/10.1037/1040-3590.19.1.44

Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., & Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Physics Reports*, *441*(1), 1–46. https://doi.org/10.1016/j.physrep.2006.12.004

Hoehl, S., & Bertenthal, B. I. (2021). *An interactionist perspective on the development of coordinated social attention*. https://doi.org/https://doi.org/10.31234/osf.io/s5tgz

Huisman, G. (2017). Social Touch Technology: A survey of haptic technology for social touch. *IEEE Transactions on Haptics*, *10*(3), 391–408. https://doi.org/10.1109/TOH.2017.2650221

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science*, *2*(4), 705–715. https://doi.org/10.1111/j.1756-8765.2010.01089.x

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), i–viii, 1–132. http://www.ncbi.nlm.nih.gov/pubmed/11428150

Jaffe, J., Stern, D. N., & Peery, J. C. (1973). "Conversational" coupling of gaze behavior in prelinguistic human development. *Journal of Psycholinguistic Research*, *2*(4), 321–329. https://doi.org/10.1007/BF01067054

Joo, H., Liu, H., Tan, L., Gui, L., Nabbe, B., Matthews, I., Kanade, T., Nobuhara, S., & Sheikh, Y.

(2015). Panoptic studio: A massively multiview system for social motion capture. *Proceedings of the IEEE International Conference on Computer Vision*, *2015 Inter*, 3334–3342. https://doi.org/10.1109/ICCV.2015.381

Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (2010). Mastering the information age solving problems with visual analytics. In *Mastering the Information Age Solving Problems with Visual Analytics*. https://doi.org/10.1016/j.procs.2011.12.035

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63. https://doi.org/10.1016/0001-6918(67)90005-4

Kim, J., Snodgrass, M., Pietrowicz, M., Karahalios, K., & Halle, J. (2013). BEDA: Visual analytics for behavioral and physiological data. *Workshop on Visual Analytics in Healthcare*, 23–27.

Kirchgässner, G., & Wolters, J. (2013). *Introduction to modern time series analysis* (2nd ed.). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73291-4

Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., & Cohen, D. (2014). Why synchrony matters during mother-child interactions: A systematic review. *PLoS ONE*, *9*(12). https://doi.org/10.1371/journal.pone.0113571

Leeds-Hurwitz, W. (1987). The social history of the natural history of an interview: A multidisciplinary investigation of social communication. *Research on Language and Social Interaction*, *20*(1–4), 1–51. https://doi.org/10.1080/08351818709389274

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(50), 13290–13295. https://doi.org/10.1073/pnas.1702493114

LePage, D. (1998). *Four- and 7-month-old infant's sensitivities to contingency during face-to-face social interaction.* [Doctoral dissertation, Concordia University]. https://spectrum.library.concordia.ca/528/

Long, B., Kachergis, G., Agrawak, K., & Frank, M. C. (2020). *Detecting social information in a dense dataset of infants' natural visual experience*. https://doi.org/10.31234/osf.io/z7tdg

Maitha, C., Goode, J. C., Maulucci, D. P., Lasassmeh, S. M. S., Yu, C., Smith, L. B., & Borjon, J. I. (2020). An open-source, wireless vest for measuring autonomic function in infants. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01394-4

Malloch, S. N., Sharp, D. B., Campbell, A. M., Campbell, D. M., & Trevarthen, C. (1997). Measuring the human voice: Analysing pitch, timing, loudness and voice quality in mother/infant communication.

Proceedings of the Institute of Acoustics, *19*(5), 495–500. http://www.co-bw.com/Music Articles Technical/Measuring the human voice Analysing pitch, timing and voice.pdf

Markova, G., Nguyen, T., & Hoehl, S. (2019). Neurobehavioral Interpersonal Synchrony in Early Development: The Role of Interactional Rhythms. *Frontiers in Psychology*, *10*(September). https://doi.org/10.3389/fpsyg.2019.02078

Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5–6), 237–329. https://doi.org/10.1016/j.physrep.2006.11.001

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters, Section A: General, Atomic and Solid State Physics*, *302*(5–6), 299–307. https://doi.org/10.1016/S0375-9601(02)01170-2

Marwan, N., Thiel, M., & Nowaczyk, N. R. (2002). Cross recurrence plot based synchronization of time series. *Nonlinear Processes in Geophysics*, *9*(3/4), 325–331. https://doi.org/10.5194/npg-9-325-2002

McDowall, J. J. (1978a). Interactional synchrony: A reappraisal. *Journal of Personality and Social Psychology*, *36*(9), 963–975. https://doi.org/10.1037/0022-3514.36.9.963

McDowall, J. J. (1978b). Microanalysis of filmed movement: The reliability of boundary detection by observers. *Environmental Psychology and Nonverbal Behavior*, *3*(2), 77–88. https://doi.org/10.1007/BF01135606

McQuown, N. A. (1971). *The natural history of an interview*. University of Chicago Library.

Mead, G. H. (1934). *Mind, self and society: from the standpoint of a social behaviourist.* University of Chicago Press.

Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychological Methods*, *23*(4), 757–773. https://doi.org/10.1037/met0000172

Nguyen, T., Abney, D. H., Salamander, D., Bertenthal, B. I., & Hoehl, S. (2021). Social touch is associated with neural but not physiological synchrony in naturalistic mother-infant interactions. *BioRxiv*. https://doi.org/10.1101/2021.01.21.427664

Nguyen, T., Bánki, A., Markova, G., & Hoehl, S. (2020). Studying parent-child interaction with hyperscanning. *Progress in Brain Research*, 1–24. https://doi.org/10.1016/bs.pbr.2020.05.003

Nguyen, T., Schleihauf, H., Kayhan, E., Matthes, D., Vrtička, P., & Hoehl, S. (2020). Neural synchrony in mother-child conversation: Exploring the role of conversation patterns. *Social Cognitive and*

Affective Neuroscience. https://doi.org/10.1093/scan/nsaa079

Paxton, A. (2015). *Coordination: Theoretical, methodological, and experimental perspectives.* [Doctoral dissertation, UC Merced]. https://escholarship.org/uc/item/5tx5s7zh

Peláez-Nogueras, M., Field, T. M., Hossain, Z., & Pickens, J. (1996). Depressed mothers' touching increases infants' positive affect and attention in still-face interactions. *Child Development*, *67*(4), 1780–1792. https://doi.org/10.1111/j.1467-8624.1996.tb01827.x

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin and Review*, *21*(1), 178–185. https://doi.org/10.3758/s13423-013-0466-4

Pereira, A. F., Yu, C., Smith, L. B., & Shen, H. (2009). A first-person perspective on a parent-child social interaction during object play. *Proceedings of the Annual Meeting of the Cognitive Science Society, 31*. https://escholarship.org/uc/item/2qf8k7h3

Pérez, S. C., & Español, S. A. (2016). Multimodal study of adult-infant interaction: A review of its origins and its current status. *Paideia*, *26*(65), 377–385. https://doi.org/10.1590/1982-43272665201613

Ragot, M., Martin, N., Em, S., Pallamin, N., & Diverrez, J. M. (2018). Emotion recognition using physiological signals: Laboratory vs. wearable sensors. *Advances in Intelligent Systems and Computing*, *608*, 15–22. https://doi.org/10.1007/978-3-319-60639-2_2

Ramírez, N. F., Hippe, D. S., & Kuhl, P. K. (2021). Comparing automatic and manual measures of parent–infant conversational turns: A word of caution. *Child Development*, 1–10. https://doi.org/10.1111/cdev.13495

Ramseyer, F. T. (2020). Motion energy analysis (MEA): A primer on the assessment of motion from video. *Journal of Counseling Psychology*, *67*(4), 536–549. https://doi.org/10.1037/cou0000407

Reidsma, D., Nijholt, A., Tschacher, W., & Ramseyer, F. (2010). Measuring multimodal synchrony for human-computer interaction. *Proceedings - 2010 International Conference on Cyberworlds, CW 2010*, *November*, 67–71. https://doi.org/10.1109/CW.2010.21

Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *NeuroImage*, *178*(May), 493–502. https://doi.org/10.1016/j.neuroimage.2018.05.060

Reindl, V., Konrad, K., Gerloff, C., Kruppa, J. A., Bell, L., & Scharke, W. (2019). Conducting hyperscanning experiments with functional near-infrared spectroscopy. *Journal of Visualized Experiments*, *2019*(143), 1–7. https://doi.org/10.3791/58807

Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.

Rogoff, B., Dahl, A., & Callanan, M. (2018). The importance of understanding children's lived experience. *Developmental Review*, *50*(March), 5–15. https://doi.org/10.1016/j.dr.2018.05.006

Ruesch, J., & Bateson, G. (1951). *Communication: The social matrix of psychiatry*. W. W. Norton & Company.

Sakurada, K., Oka, A., Kiso, R., Shimabukuro, L., Ueno, A., & Nakatani, M. (2019). *Baby touch: quantifying visual-haptic exploratory behaviors in infants of sensory-motor development* (pp. 241–246). https://doi.org/10.1007/978-981-13-3194-7_55

Scheflen, A. E. (1965). Quasi-courtship behavior in psychotherapy. *Psychiatry*, *28*(April), 245–257. https://doi.org/10.1080/00332747.1965.11023433

Scheflen, A. E. (1967). On the structuring of human communication. *American Behavioral Scientist*, *10*(8), 8–12. https://doi.org/10.1177/0002764201000803

Schroer, S. E., Smith, L. B., & Yu, C. (2019). Examining the multimodal effects of parent speech in parent-infant interactions. *CogSci*, 1015–1021.

Serra, J., Miguel, H., Moura, A. A., Sampaio, A., & Pereira, A. F. (2020). The effect of play task on maternal touch patterns when interacting with their 12 months-old infants: An exploratory study. *Infant Behavior and Development*, *59*(June 2019), 101438. https://doi.org/10.1016/j.infbeh.2020.101438

Shockley, K., Butwill, M., Zbilut, J. P., & Webber, C. L. (2002). Cross recurrence quantification of coupled oscillators. *Physics Letters, Section A: General, Atomic and Solid State Physics*, *305*(1–2), 59–69. https://doi.org/10.1016/S0375-9601(02)01411-1

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*(4), 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science*, *14*(1), 9–17. https://doi.org/10.1111/j.1467-7687.2009.00947.x

Smith, L. B., Yu, C., Yoshida, H., & Fausey, C. M. (2015). Contributions of head-mounted cameras to studying the visual environments of infants and young children. *Journal of Cognition and Development*, *16*(3), 407–419. https://doi.org/10.1080/15248372.2014.933430

Sousa, E., Malheiro, T., Bicho, E., Erlhagen, W., Santos, J., & Pereira, A. F. (2016). *MUVTIME: A multivariate time series visualizer for behavioral science*. *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - IVAPP,* *(VISIGRAPP 2016)*, *March*, 165–176. https://doi.org/10.5220/0005725301650176

Stack, D. M. (2004). Touching during mother-infant interactions. In T. Field (Ed.), *Touch and Massage in Early Child Development* (pp. 49–81). Johnson & Johnson Pediatric Institute, L.L.C.

Stack, D. M., & LePage, D. E. (1996). Infants' sensitivity to manipulations of maternal touch during face-to-face interactions. *Social Development*, *5*(1), 41–55. https://doi.org/10.1111/j.1467-9507.1996.tb00071.x

Stack, D. M., LePage, D. E., Hains, S., & Muir, D. W. (1996). Qualitatitive changes in maternal touch as a function of instructional condition during face-to-face social interactions. *Infant Behavior and Development*, *19*(3), 761. https://doi.org/10.1016/S0163-6383(96)90815-2

Stack, D. M., & Muir, D. W. (1990). Tactile stimulation as a component of social interchange: New interpretations for the still-face effect. *British Journal of Developmental Psychology*, *8*(2), 131–145. https://doi.org/10.1111/j.2044-835x.1990.tb00828.x

Stack, D. M., & Muir, D. W. (1992). Adult tactile stimulation during face-to-face interactions modulates five-month-olds' affect and attention. *Child Development*, *63*(6), 1509–1525. https://doi.org/10.1111/j.1467-8624.1992.tb01711.x

Stepakoff, S. (1999). *Mother-infant tactile communication at four months: Effects of infant gender, maternal ethnicity, and maternal depression.* St. John's University.

Stern, D. N. (1971). A micro-analysis of mother-infant interaction: Behavior regulating social contact between a mother and her 3 1/2-month-old twins. *Journal of the American Academy of Child Psychiatry*, *10*(3), 501–517. https://doi.org/10.1016/S0002-7138(09)61752-0

Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver.* Wiley-Interscience.

Stern, D. N. (2002). *The first relationship: infant and mother* (2nd ed.). Harvard University Press.

Stern, D. N., Beebe, B., Jaffe, J., & Bennett, S. L. (1977). The infant's stimulus world during social interaction. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction: Proceedings of the Loch Lomond Symposium* (pp. 177–202). Academic Press.

Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: Two modes of communication within the mother-infant dyad. *Annals of the New York Academy of Sciences*, *263*(1), 89–100. https://doi.org/10.1111/j.1749-6632.1975.tb41574.x

Stern, D. N., Spieker, S., Barnett, R., & Mackain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, *10*(1), 1–15.

https://doi.org/10.1017/S0305000900005092

Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, *18*(5), 727–735. https://doi.org/10.1037/0012-1649.18.5.727

Sullivan, H. S. (1955). *The interpersonal theory of psychiatry*. Routledge.

Sun, F. T., Miller, L. M., & D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *NeuroImage*, *21*(2), 647–658. https://doi.org/10.1016/j.neuroimage.2003.09.056

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, *20*(6). https://doi.org/10.1111/desc.12456

Tragger, G. L. (1958). Paralanguage: a first approximation. *Studies in Linguistics*, *13*, 1–11.

Tronick, E. Z. (1995). Touch in mother-infant interaction. In T. M. Field (Ed.), *Touch in Early Development* (pp. 53–65). Lawrence Erlbaum Associates.

Tronick, E. Z., Adamson, L., Als, H., & Brazelton, T. B. (1975). Infant emotions in normal and pertubated interactions. *Paper Presented at the Biennial Meeting of the Society for Research in Child Development*.

Tronick, E. Z., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Tronick, E. Z., Als, H., & Brazelton, T. B. (1980). Monadic phases: A structural descriptive analysis of infant-mother face to face interaction. *Merrill Palmer Quarterly*, *26*(1), 3–24. https://www.jstor.org/stable/23084074

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, *9*, 151–176. https://doi.org/10.1146/annurev-clinpsy-050212-185510

Tschacher, W., & Haken, H. (2019). *The process of psychotherapy: Causation and chance*. Springer International Publishing. https://doi.org/10.1007/978-3-030-12748-0

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*(NOV), 1–13. https://doi.org/10.3389/fpsyg.2014.01323

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2). Addison-Wesley.

Wass, S. V., Whitehorn, M., Marriott Haresign, I., Phillips, E., & Leong, V. (2020). Interpersonal neural entrainment during early social interaction. *Trends in Cognitive Sciences*, *24*(4), 329–342.

https://doi.org/10.1016/j.tics.2020.01.006

Weakland, J. H. (1967). Communication and behavior - an introduction. *American Behavioral Scientist*, *10*(8), 1–4. https://doi.org/10.1177/0002764201000801

Xu, T. L., de Barbaro, K., Abney, D. H., & Cox, R. F. A. (2020). Finding structure in time: Visualizing and analyzing behavioral time series. *Frontiers in Psychology*, *11*, 1–22. https://doi.org/10.3389/fpsyg.2020.01457

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy*, *13*(3), 229–248. https://doi.org/10.1080/15250000802004437

Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Current Biology*, *26*(9), 1235–1240. https://doi.org/10.1016/j.cub.2016.03.026

Yu, C., Smith, L. B., & Pereira, A. F. (2008). Embodied solution: The world from a toddler's point of view. *2008 IEEE 7th International Conference on Development and Learning, ICDL*, *June 2014*, 97–102. https://doi.org/10.1109/DEVLRN.2008.4640812

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, *22*(1), 1–12. https://doi.org/10.1111/desc.12735

Yu, C., Yurovsky, D., & Xu, T. L. (2012). Visual data mining: An exploratory approach to analyzing temporal patterns of eye movements. *Infancy*, *17*(1), 33–60. https://doi.org/10.1111/j.1532-7078.2011.00095.x

CHAPTER V – GENERAL DISCUSSION

General Discussion

In Chapter I of this dissertation, we have introduced conversational turn-taking as a form of interpersonal coordination. We proposed to study the development of the temporal properties of turn-taking, by using turn-transition duration as a measure that effectively captures the temporal coordination necessary for the minimal-gap minimal-overlap standard of adult conversations. We have explored the most prominent models of turn-taking, some as the dialogical approach (Cassotta et al., 1964; Jaffe et al., 2001; Jaffe & Feldstein, 1970) and the interaction engine hypothesis (Levinson, 2006, 2019), with methodological and theoretical implications for our work. We proceeded to explore evolutionary, cross-cultural and developmental evidence of the foundations of turn-taking. To finally explore the contributions of the most recent studies in the developmental trajectory of turn-transition duration, and the challenges the challenges that emerge from this research.

In Chapter II, we have presented the first empirical study (Study 1) of this dissertation, where we investigate the developmental trajectory of gap and overlap durations, from 7 to 12 months, and the hypothetical effect of object-oriented play in turn-transition duration. We used a longitudinal design, with two age points (7, 12 months), and measured gap and overlap durations in 25 mother-infant dyads in three tasks: (1) free-play with toys, (2) free-play without toys, and (3) challenging object-play. The effects of task, infant's age and direction of turn-transition (mother to infant, infant to mother) were analyzed (Lourenço, Pereira, et al., 2021).

We have demonstrated that there is a significant difference in infant's gap durations, when comparing face-to-face and object-oriented mother-infant interactions. When objects are removed from the interaction, infants exhibit much shorter median gap durations, which became even shorter from 7 to 12 months (Lourenço, Pereira, et al., 2021).

In Chapter III, we have reported the second empirical study (Study 2) of this dissertation, where we examine the developmental trajectories of gaps and overlaps, in childhood, and compare them to overall turn-transition duration. For that, we examined 44 mother-child free-play interactions, in a cross-sectional design, with 3- to 5-year-old children. We measured gap and overlap duration has independent and aggregated metrics – floor-transfer offset (FTO). The effects of children's age and direction of turn-transition (mother to child, child to mother) were analyzed independently for each metric – gap, overlap, FTO (Lourenço et al., 2022).

We have shown that FTO durations in 3- to 5-years-old children are significantly longer than their mothers, and still longer than the FTO durations expected in adults. However, the contribution of gap and overlap duration for that outcome its considerably different and indicative that, as in infancy, gaps and overlaps have different developmental trajectories throughout childhood. Gap durations of 3- to 5-years-old children, were still slightly (but not significantly) increasing; and were significantly longer than their mothers' gap durations; and far from the minimal-gap standard of adult conversation. Meanwhile, overlap durations were considerably shorter than gaps, and shorter than the durations in infancy (Hilbrink et al., 2015; Lourenço, Pereira, et al., 2021). Furthermore, their duration were not significantly different than their mothers; and comparable to the minimal-overlap standard of adult conversation (Lourenço et al., 2022).

Results also provided a good indication of the informative value of our methodological options. Measuring gaps and overlaps, as aggregated (FTO) and independent metrics, shown the differences, in developmental trajectory and in the contribution to children's turn-transition timing, between gaps and overlaps. Examining turn-transition timing in all types of vocalizations, instead of focusing on specific question-response pairs, provided us with a more comprehensive reference for children's general turn-transition timing. One that shares the same vocalization criteria, in childhood and in infancy (Lourenço, Pereira, et al., 2021), and, for that, can be immediately comparable, when examining the developmental continuities between preverbal and verbal children (Lourenço et al., 2022).

In Chapter IV of this dissertation we presented a methodological study (Study 3), that is an overview of the microanalysis, as a methodological approach. We explored the historical and theoretical origins of microanalysis as a method for observing and coding interactional behavior. We examined how microanalysis transformed our understanding of the structural and temporal aspects of mother-infant interactions. We described the methodological challenges to the application of a completely microanalytic approach and reviewed some of the emergent technological innovations and techniques that may improve our ability to capture, explore and model the multidimensionality of interactive behavior. We illustrated the transformative potential of these innovations with the use cases of egocentric vision research and daily activity sensing research. Finally, we discussed how these innovations can help us go beyond the limitations and obstacles found in classical microanalytic work and reach the full potential of this methodological approach, as it was originally envisioned (Lourenço, Coutinho, et al., 2021).

Throughout this dissertation, we have centered on two dimensions of turn-taking research: empirical and methodological. On one side, we followed the intricate developmental thread of turn-taking, by examining how one of its essential properties - minimal-gap minimal-overlap – develops over time. On the other, we examined the implications of certain methodological choices in our pursuit of understanding about the development of turn-taking, interpersonal coordination and interactional behavior, in general. These are by no means independent dimensions of research. As we have seen with the birth of microanalysis and its application to developmental research. Methodological aspects guide our process of research to the very ability to observe our object of study (Stern, 2002). Their empirical application and results, in turn, provide not only the necessary data to improve our understanding of a phenomenon, but also invaluable information about the potential and limitations of a methodological framework.

Turn-taking is an essential aspect of human conversation, and possibly an element of communication and social interaction beyond our species (Levinson, 2019; Pika et al., 2018). What we presented in this dissertation, was a microanalytic perspective to the study of turn-taking that attempts to reconcile these different levels of abstraction – conversation, communication, interaction – while still considering the linguistic processing challenges that have emerged from psycholinguistic research.

In Chapter I of this dissertation, we explored how this more psycholinguistic perspective followed from the cognitive implications of a conversational analysis approach to turn-taking (Hilbrink et al., 2015; Sacks et al., 1974). Curiously though, when conversation analysis developed and start systematizing the organization of turn-taking (Sacks et al., 1974), it shared many similarities with the microanalytic approach. Building on the same natural history insights, use of recording technology, and detailed analysis that characterized microanalytic work (McQuown, 1971). Ultimately, with similar goals: study conversation as it naturally occurs, with an appreciation for nonverbal behavior (Lynch & Bogen, 1994).

Through the lens of the interactive engine hypothesis (Levinson, 2006, 2019), the more psycholinguistic perspective, has been expanding our understanding of turn-taking development, in the last decade. By demonstrating infants' sensitivity to turn-timing and their predictive ability (Casillas & Frank, 2013, 2017; Thorgrimsson, 2014). By displaying the early differentiation between gaps and overlaps (Hilbrink et al., 2015). And by enlightening that different levels of linguistic complexity can contribute to a non-linear and slow developmental progression of turn-transition timing towards adult standards (Casillas et al., 2016; Lindsay et al., 2019; Stivers et al., 2018).

In Chapter I, we have also critically analyze some of the challenges that have emerged from that work. How an increase in gap duration, from 5 to 9 months, has been interpreted as a confirmation of a theoretically expected slowdown in infant's turn-transition timing, due to the integration between early turn-taking skills and additional linguistic processing (Hilbrink et al., 2015). Even though infant's communicative abilities throughout that period are still mostly nonverbal, and the slowdown was only predicted from 12 months onwards. Inversely, how the subsequent (9 to 18 months) decrease in gap duration is interpreted as an indicator of progressive integration between turn-taking and language processing (Hilbrink et al., 2015). Even though infants are still in the earlier stages of verbal communication, and, as later confirmed, this progression is mostly likely slow and non-linear (Casillas et al., 2016; Lindsay et al., 2019; Stivers et al., 2018). How the study of turn-transition duration in verbal children did not cover a larger spectrum of vocalizations, privileging specific semantic contingent exchanges: question-response pairs (Casillas et al., 2016; Stivers et al., 2018). Or how, in verbal children, overlaps are assumed, as a measure of turn anticipation, as negative gaps; and therefore, aggregated metrics – such as response latency – of overlap and gap duration are the standard (Casillas et al., 2016). While the differences in the developmental trajectory of gap and overlap, and their individual contribution to children's turn-timing may be overlooked.

In this dissertation, we have taken a more microanalytic approach to the study of turn-taking (Lourenço, Coutinho, et al., 2021). In both empirical studies (Lourenço et al., 2022; Lourenço, Pereira, et al., 2021) we used the same binary logic of the AVTA model (Cassotta et al., 1964) and the turn-allocation rule specified by the dialogical systems (Jaffe & Feldstein, 1970) approach to derive dyadic states of turn-taking from the whole vocalizations exchanged between interactional partners (Jaffe et al., 2001). This ensured that each partner's vocalizations contributed to the very definition of what constituted a turn-transition – defined by the continuous relation between the vocal behaviors of each partner, and not by their semantic contingency. For that, it has also guaranteed that the turn-transitions analyzed in our work were not a sliced perspective on vocal turn-taking – based solely on semantic pair adjacent transitions (e.g., question-response pairs) –, but instead a wider representation of the spectrum of the (verbal and nonverbal) vocalizations used for turn-transitioning. Using the same coding strategy and comparable metrics of turn-transition durations, also helped highlight the developmental continuities in infant (Lourenço, Pereira, et al., 2021) and children turn-taking (Lourenço et al., 2022).

There were nonetheless limitations to our approach. More generally, we measured temporal coordination only in terms of turn-transition duration. As we have seen in Study 3, there are other alternatives that could help us explore vocal coordination by considering the vocalization time-series of each interactional partner. This would facilitate the addition of other nonverbal channels of behavior (Lourenço, Coutinho, et al., 2021). We believe nonetheless, that the strategy we adopted was the best to the research questions we intended to answer.

Specifically to the design of Study 1 (Lourenço, Pereira, et al., 2021), we wanted to understand the differences in turn-transition duration, between face-to-face and object-oriented interactions in the second half of infant's first-year. Specifically, in relation to an hypothetical increase in gap duration between 5 and 9 months (Hilbrink et al., 2015). Nevertheless, we were working with a database that only had longitudinal data at 7 and 12 months to test our hypothesis. We cannot discard that an increase at 9 months, between the age points we have analyzed, can actually occur. we cannot discard that an increase at 9 months – between the age points we have analyzed – can actual occur. Only that at 7 and 12 months, infants can have much shorter median gap durations than previously reported (Hilbrink et al., 2015), which are highly dependent on the type of interaction (object-oriented or face-to-face). Although the absence of a significant difference in gap duration, between 9 and 12 months, in that study (Hilbrink et al., 2015), helps our argument.

An additional limitation in Study 1 (Lourenço, Pereira, et al., 2021), was the absence of counter-balancing between face-to-face and object-oriented conditions. The conditions in the database used in our study, followed a fixed order that has been applied to the study of mother's sensitivity (Mateus et al., 2021). The order progresses from a less challenging condition to the mother, where play is less-structured with familiar objects (free-play with toys); to a more challenging condition to the mother, where play depends on mother's ability to engage the infant in face-to-face interaction without toys as a resource (free-play without toys); to the most challenging condition, where mothers help their infants play with toys that are slightly above their developmental level of their infants (challenging-object play).

We used the free-play with toys condition and the challenging-object play condition as instances of object-oriented interaction; and the free-play without toys condition as the instance of face-to-face interaction. Given that the face-to-face segment of the interaction followed an object-oriented segment and was proceeded by another object-oriented segment, we had A-B-A order that could provide an approximation to the differences between transitioning from an object-oriented to a face-to-face interaction, and transitioning from a face-to-face to an object-oriented interaction. The results were reassuring since turn-transition durations were similar in the two object-oriented tasks, and gap duration was significantly longer for both tasks than in the face-to-face interaction (Lourenço, Pereira, et al., 2021).

Regarding Study 2 (Lourenço et al., 2022) we believe that if we had used a longitudinal design instead, it may had facilitated age comparisons. Additionally, if we had completed the analysis of Study 1, before designing and implementing Study 2, we would have included a comparison between an object-oriented and a face-to-face condition, in the design. Initially, we had no reason to believe that 3- to 5-years-olds would have similar difficulties in coordinating object-play with the flow of turn-taking, than those hypothesized for 7 and 12 months infants.

Finally, a potential limitation of Study 3 (Lourenço, Coutinho, et al., 2021) is that we had to be selective in our account of events, theories, studies, technologies and techniques. It was out of the scope of the study to be an history of science, or an exhaustive methodological and technological review. We sought to provide the necessary ingredients to the advancement of microanalytic research and the developmental study of interactive behavior. Perhaps, one other valid limitation in Study 3, is that, indeed, it offers only the ingredients, and not the recipes.

Future Research Directions

We would like to highlight three threads of research that follow from the results and arguments presented in this dissertation: (1) the developmental trajectories of gap and overlap duration, in the transitional period between late infancy and early childhood; (2) the progression of the interference effect of object-oriented interaction in turn-transition timing; and (3) turn-taking across other channels of interactive behavior and in relation to vocalizations. We will continue by discussing this three threads, integrating our conclusions with formulations that will guide our future research.

The Developmental Trajectories of Gap and Overlap Timing: From 12 to 36 Months

Comparing the results of 7- and 12-months-old infants (Lourenço, Pereira, et al., 2021) to the results of 3- to 5-years-old children (Lourenço et al., 2022), we have suggested that gap and overlap timing may be developing in different directions, throughout the transitional period between 12 and 36 months. While overlap duration may be slightly decreasing towards the minimal-overlap standard of adult turn-taking; gap duration, in general – and not considering the differences between levels of linguistic complexity –, may be getting slightly longer and further from the adult minimal-gap standard.

When discussing the differences between gaps and overlaps (Chapter I) and the results in Study 2 (Lourenço et al., 2022), we have considered that if, indeed, overlap duration is converging to the minimal-overlap standard, this should be noticeable in overall turn-transition (response latency/FTO) durations, throughout this period (12-36 months). If response latencies are getting longer (Lourenço et al., submitted), or at least maintaining similarly high (Casillas et al., 2016), this may be attributable to an hypothetical increase in gap duration, from 12 to 36 months. This hypothesis would, indeed, be coherent with meta-analytic data indicating that gap duration may be increasing at least up until 40 months (Nguyen et al., 2022). We also suggest, that an actual increase in gap duration throughout this period, which is characterized by a large spectrum of linguistic developments (Casillas et al., 2016), may in fact be more revealing of the hypothetical slowdown that is expected when early turn-taking abilities and linguistic skills begin to be integrated (Hilbrink et al., 2015).

To examine the developmental trajectory of gap and overlap duration throughout this period, we plan to follow up our research with a new longitudinal study, or series of longitudinal studies, that will encompass the period between 12 and 36 months, and analyze gap and overlap duration, both as independent and aggregated (FTO) measures, in all temporally-contingent vocalizations between toddlers and their caregivers. We would prioritize data collection at equidistant age points 12, 18, 24, 30 and 36 months, since they are also turning-points for many of the developmental milestones throughout this period.

The Interference Effect of Object-oriented Interactions in Turn-transition Duration

In Study 1 (Lourenço, Pereira, et al., 2021) we have demonstrated that, at least at 7 and 12 months, infant's gap duration could be significantly reduced by removing object-play from mother-infant interaction. We interpreted these results as an hypothetical interference in the flow of turn-taking, due to the immaturity of the motor skills used in object-oriented interactions. When considering the potential effect of object-play in turn-transition timing (Chapter I, Chapter II), we have pointed that throughout this period (7-12 months), infants develop basic locomotion, body-weight support and balance, and object-manipulation skills that encourage environmental exploration and agency in object-play (Adolph & Hoch, 2019). However, these skills are still too rudimental and may involve an additional effort from the infant, that may compete with maintaining appropriate timing in their interactional vocal exchanges.

We predict that this hypothetical interference may weaken and became less evident, as motor skills that enable object-oriented interactions mature to a point that children can easily coordinate object-manipulation with parallel conversation. Indeed, we have not found any indications, in Study 2 (Lourenço et al., 2022) that 3-years-olds and older children's turn-transition timing may be affected in the same way when engaged in object-oriented interactions. In fact, their FTO turn-transition duration was faster than the response latency for the question-response pairs of same age children in other studies (Casillas et al., 2016; Stivers et al., 2018).

To investigate if similar differences between face-to-face and object-oriented interactions can be found beyond 12 months old, we could add a free object-play condition to the longitudinal study we have suggested above, between 12 and 36 months. To test the hypothesis that a potential smoothing of that difference may be due to the maturation of motor skills, we would assess toddler's motor development using Griffiths Scales of Child Development III (Stroud et al., 2016).

Complementarily, in order to gain a better insight on the mechanics by which object-oriented interactions may affect turn-transition timing – and also help discard alternative explanations –, we could use a different microanalytic approach to data collection and coding of object-oriented behavior. To minimize the obstruction of relevant visual information while recording interaction, we could use several cameras – including head-mounted cameras, to capture first-person perspectives (Pereira et al., 2014; Yoshida & Smith, 2008) – to follow toddler's and mother's object-oriented behaviors (e.g., gaze direction, grasping, object-manipulation). Perhaps the use of a computer vision approach (e.g., OpenPose; Cao et al., 2021), could help in the automatic detection of object-oriented behaviors. Based on the binary logic of the AVTA model (Cassotta et al., 1964), we could further attempt to cross the several recorded perspectives to derive the presence or absence of specific object-oriented behaviors, and explore dyadic states of joint-attention and joint-action.

Additionally, we have considered that the interference of object-oriented interaction in turn-transition timing can expand beyond motor skill, if conceived as a competition between infant's resources to engage in object-oriented behavior and the resources to maintain the flow of vocal exchanges. We conjecture that a similar competition may occur – with different resources, and at other stages of development - when challenging object-oriented tasks require additional resources (e.g., cognitive).

Using the same 3- to 5-years-old cross-sectional sample of Study 2 (Lourenço et al., 2022), we tested two additional conditions of mother-child object-oriented interactions that should elicit different cognitive skills. On one task, we asked dyads to use colored building-blocks to jointly assemble several figures depicted in images showing only the final result. On the other task, we asked dyads to co-create a story based on story sequencing cards. We have began the process of segmentation and coding of the data collected. Once manual segmentation of vocalizations is over and the automated coding of turn-taking states complete we will compare the data between all three conditions: free-play, building blocks, story sequences.

Turn-taking Beyond Vocalization

Most of this dissertation has been dedicated to the study of vocal turn-taking. We have differentiated from other more psycholinguistic approaches (Casillas et al., 2016; Stivers et al., 2018) by measuring a larger spectrum of vocalizations, including both verbal and nonverbal vocal exchanges. But conversational turn-taking expands to other nonverbal channels of communication.

In Chapter I, when discussing evolutionary evidence of the foundations of turn-taking, we have seen that great apes use instrumental gestures in their social exchanges with an alternation pattern timed similarly to human conversational turn-taking (Rossano, 2013; Rossano & Liebal, 2014). In Study 3 (Lourenço, Coutinho, et al., 2021), when discussing the contributions of microanalysis to the understanding of the temporal structure of mother-infant interactions, we have seen that both the alternation pattern of turn-taking and simultaneous patterns of temporal coordination can be found in the multimodal stream of interactional behaviors (Fogel, 1977; Stern et al., 1975).

Several nonverbal behaviors can be identified as particularly relevant to conversational turn-taking (e.g, posture, eye gaze, facial expression, head movement, manual gesture, or touch). These may be used to signal selection at any point of the turn (e.g., eye gaze); as a non-vocal response (e.g., head movement, manual gesture); or as a double response when movement and vocalization are used in the same response (Clark & Lindsey, 2015; Stivers et al., 2009).

We know that from, at least 10 months onwards, gaze and manual gesture, particularly pointing and reaching, are used by young children in substitution or in addition to speech (Clark & Lindsey, 2015). And in adults, head movement (e.g., nods, shrugs, shakes) and gaze explain cultural variation in turn-transition timing (Stivers et al., 2009).

Few studies have employed the same framework used to measure vocal turn-taking, in terms of turn-transition (gap and overlap) durations, to other nonverbal channels. In a longitudinal case-study, from 19 to 40 months, using the same approach to measure vocal and gestural turn-transition durations, Clark & Lindsay (Clark & Lindsey, 2015) shown that gestural response latencies to Where and Which questions were overall faster than vocalizations, and that in double responses, where gestures and vocalization are displayed, gestures were produced faster.

Pereira et al. (2008) have used instead a time-series approach to study the temporal dynamics of head and hand movements in relation to parent-child interpersonal coordination, while learning new words. A partner’s head movements were inversely related to the other’s hand movements, indicative of a turn-taking pattern between the two channels. The body coupling dynamics between partners was also predictive of infant's word learning.

Expanding on the suggestion we delineated above to research object-oriented behavior in a longitudinal study of the developmental trajectory of turn-transition duration, from 12 to 36 months, we could use the same strategy to code head movement and manual gestures that may be used instead of or in addition to vocalization. And analyze them, as in Clark and Lindsay (Clark & Lindsey, 2015), in terms of turn-transition duration, as a vocal response.

Perhaps sooner, we will be able to expand the analysis of the mother-infant interactions examined in Study 1 (Lourenço, Pereira, et al., 2021), with the addition of social touch behavior. By crossing our data on vocal coordination with the data of a parallel project, sharing the same database, that uses a classical microanalytical approach to examine mother's touch behavior (Serra et al., 2020), we will be able to explore the temporal relations between both channels of behavior.

On another front, we tested, in the database we collected for Study 2 (Lourenço et al., 2022), the use of inertial motion sensors to track arm and head movement. Once the additional experimental conditions (building-blocks, story sequence) are also segmented and coded, we will be able to explore the potential relations between movement, type of object-oriented task, and vocalizations. We will consider instead a multivariate time series approach to the analysis. Rather than coding dyadic states of turn-taking in a univariate measure, we will use the raw vocalizations of each interactional partner as two time series that can be temporally related to the time series generated from each partner's arm and head movement.

In Study 3 (Lourenço, Coutinho, et al., 2021), we have described some alternatives to explore and model such temporal relations between multivariate time series. It is important to understand that, as a multidimensional phenomenon, interpersonal coordination between different streams of behavior may not be observable between all streams, in the same direction, or on the same temporal scale.

General Conclusion

Turn-taking is an essential part of human communicative behavior. It is a natural and effortless process that even young infants can partake, but has a slow progressing and non-linear developmental trajectory, before completely converging to adult standards. As an interactional behavior phenomenon, it is also multidimensional and surprisingly complex to investigate. We believe that the contributions of this dissertation, to the developmental study of turn-taking, bring new and pertinent information to the research field. Not only because they expand our knowledge about the developmental trajectory of turn-taking, but also because they promote a better understanding of the implications of some methodological options we take in the pursuit for knowledge. There is still much to understand about the development of turn-taking. Finer-grained analysis of the temporal properties of turn-transition, with longitudinal data and larger samples are just emerging; and the focus on nonverbal behavior is still residual. We hope that the contributions in this dissertation can impact the research field as deeply as they will certainly impact our future research, and motivate further research on the temporal microdynamics of interpersonal coordination.

References

Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, *70*(September 2018), 141–164. https://doi.org/10.1146/annurev-psych-010418-102836

Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2021). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(1), 172–186. https://doi.org/10.1109/TPAMI.2019.2929257

Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn-taking, timing, and planning in early language acquisition. *Journal of Child Language*, *43*(6), 1310–1337. https://doi.org/10.1017/S0305000915000689

Casillas, M., & Frank, M. (2013). The development of predictive processes in children's discourse understanding. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 299–304. http://langcog.stanford.edu/papers/CF-cogsci2013.pdf

Casillas, M., & Frank, M. C. (2017). The development of children's ability to track and predict turn structure in conversation. *Journal of Memory and Language*, *92*, 234–253. https://doi.org/10.1016/j.jml.2016.06.013

Cassotta, L., Feldstein, S., & Jaffe, J. (1964). AVTA: A device for automatic vocal transiction analysis. *Journal of the Experimental Analysis of Behavior*, *7*(1), 99–104. https://doi.org/10.1901/jeab.1964.7-99

Clark, E. V., & Lindsey, K. L. (2015). Turn-taking: A case study of early gesture and word use in answering WHERE and WHICH questions. *Frontiers in Psychology*, *6*(July), 1–7. https://doi.org/10.3389/fpsyg.2015.00890

Fogel, A. (1977). Temporal organization in mother-infant face-to-face interaction. *Studies in Mother-Infant Interaction*, *January 1977*, 119–152.

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology*, *6*(September), 1–12. https://doi.org/10.3389/fpsyg.2015.01492

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child*

Development, *66*(2), i–viii, 1–132. http://www.ncbi.nlm.nih.gov/pubmed/11428150

Jaffe, J., & Feldstein, S. (1970). *Rhythms of dialogue*. Academic Press.

Levinson, S. C. (2006). On the human interactional engine. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality. Culture, Cognition and Interaction* (pp. 39–69). Berg.

Levinson, S. C. (2019). Interactional foundations of language: The interaction engine hypothesis. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 189–200). MIT Press.

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychological Science*, *30*(4), 504–515. https://doi.org/10.1177/0956797618822802

Lourenço, V., Coutinho, J., & Pereira, A. F. (2021). Advances in microanalysis: Magnifying the social microscope on mother-infant interactions. *Infant Behavior and Development*, *64*(April), 101571. https://doi.org/10.1016/j.infbeh.2021.101571

Lourenço, V., Pereira, A. F., Sampaio, A., & Coutinho, J. (2021). Turn-taking in object-oriented and face-to-face interactions: A longitudinal study at 7 and 12 months. *Psychology & Neuroscience*. Advance online publication. https://doi.org/10.1037/pne0000276

Lourenço, V., Serra, J., Coutinho, J., & Pereira, A. F. (2022). *Turn-taking in free-play interactions: A Cross-sectional study from 3 to 5 years*. [Manuscript submitted for publication]. School of Psychology, University of Minho.

Lynch, M., & Bogen, D. (1994). Harvey Sacks's primitive natural science. *Theory, Culture & Society*, *11*(4), 65–104. https://doi.org/10.1177/026327694011004003

Mateus, V., Osório, A., Miguel, H. O., Cruz, S., & Sampaio, A. (2021). Maternal sensitivity and infant neural response to touch: An fNIRS study. *Social Cognitive and Affective Neuroscience*, *16*(12), 1256–1263. https://doi.org/10.1093/scan/nsab069

McQuown, N. A. (1971). *The natural history of an interview*. University of Chicago Library.

Nguyen, V., Versyp, O., Cox, C., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the development of turn taking in adult–child vocal interactions. *Child Development*, *4*, 1–42. https://doi.org/10.1111/cdev.13754

Pereira, A. F., Smith, L. B., & Yu, C. (2008). Social coordination in toddler's word learning: Interacting systems of perception and action. *Connection Science*, *20*(2–3), 73–89. https://doi.org/10.1080/09540090802091891

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin and Review*, *21*(1), 178–185. https://doi.org/10.3758/s13423-013-0466-4

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: Bridging the gap between human and animal communication. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880). https://doi.org/10.1098/rspb.2018.0598

Rossano, F. (2013). Sequence organization and timing of bonobo mother-infant interactions. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, *14*(2), 160–189. https://doi.org/10.1075/is.14.2.02ros

Rossano, F., & Liebal, K. (2014). *"Requests" and "offers" in orangutans and human infants*. August*, 335–364. https://doi.org/10.1075/slsi.26.13ros

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Serra, J., Miguel, H., Moura, A. A., Sampaio, A., & Pereira, A. F. (2020). The effect of play task on maternal touch patterns when interacting with their 12 months-old infants: An exploratory study. *Infant Behavior and Development*, *59*(March), 101438. https://doi.org/10.1016/j.infbeh.2020.101438

Stern, D. N. (2002). *The first relationship: Infant and mother* (2nd ed.). Harvard University Press.

Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: Two modes of communication within the mother-infant Dyad. *Annals of the New York Academy of Sciences*, *263*(1), 89–100. https://doi.org/10.1111/j.1749-6632.1975.tb41574.x

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stivers, T., Sidnell, J., & Bergen, C. (2018). Children's responses to questions in peer interaction: A window into the ontogenesis of interactional competence. *Journal of Pragmatics*, *124*, 14–30. https://doi.org/10.1016/j.pragma.2017.11.013

Stroud, L., Foxcrot, C., Green, E., Bloomfield, S., Cronje, J., Hurter, K., ..., & Venter, D. (2016). *Griffiths scales of child development 3rd Ed. Part I: Overview, development and psychometric properties*. Hogreffe.

Thorgrimsson, G. B. (2014). *Infants' understanding of communication as participants and observers*. [Unpublished doctoral dissertation]. Radbound University Nijmegen.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy*, *13*(3), 229–248. https://doi.org/10.1080/15250000802004437