Cognition 239 (2023) 105568

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Turn-taking in free-play interactions: A cross-sectional study from 3 to 5 years

Vladimiro Lourenço [a,*], Juliana Serra [a], Joana Coutinho [a], Alfredo F. Pereira [a,b,*]

[a] Psychology Research Center (CIPsi), School of Psychology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[b] NOVA School of Science and Technology, Center of Technology and Systems (UNINOVA-CTS), NOVA University Lisbon, 2829-516 Lisbon, Portugal

ARTICLE INFO

Keywords:
Turn-taking
Mother-child interaction
Floor-transfer offset
Gaps
Overlaps

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 modeling 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. Specifically, mother-to-child turn-transitions produced significantly longer FTO and gap durations than child-to-mother turn-transitions, 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 of age. 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.

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, Schegloff, & Jefferson, 1974). Turn-transition is tightly synchronized: partners exchange the floor by either leaving a small silence (a gap) or a small period of simultaneous speed (an overlap). This temporal coordination, described as the minimal-gap minimal-overlap phenomenon (Levinson & Torreira, 2015), acts as a glue that holds turns together (Casillas, 2014).

1. 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 which overlaps assume negative latency values, and gaps positive values. In adult-adult conversations, gap durations averages around 200 ms, and FTO has an unimodal and slightly asymmetrical distribution, similar across different kinds of cultures that also differ in language structure (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 (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 one of its components, and on which languages builds up. We will refer to the processes that subserve social coordination, in particular turn-

* Corresponding authors at: Psychology Research Center (CIPsi), School of Psychology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal.
E-mail addresses: id6774@uminho.pt (V. Lourenço), juliana.serra@psi.uminho.pt (J. Serra), joanacoutinho@psi.uminho.pt (J. Coutinho), ap@uninova.pt (A.F. Pereira).

https://doi.org/10.1016/j.cognition.2023.105568
Received 19 May 2022; Received in revised form 29 May 2023; Accepted 16 July 2023
Available online 26 July 2023

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, Gattis, & Levinson, 2015; Levinson, 2019).

1.1. Turn-taking development

Observational studies using microanalysis (Lourenço, Coutinho, & Pereira, 2021) have shown, that infants are able to engage in contingent interaction across different modalities of communication (Bornstein, Putnick, Cote, Haynes, & Suwalsky, 2015; Brazelton, Koslwski, & Main, 1974; Fogel, 1977; Kaye & Fogel, 1980; Stern, 1971; Tronick, Als, & Brazelton, 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 infants are verbal (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, Als, Adamson, Wise, & Brazelton, 1978) –, have shown indeed that infants, by at least 3 months of age, are sensitive to alterations in contingency (Murray & Trevarthen, 1985; Striano, Henning, & Stahl, 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, Alson, Jaffe, Feldstein, & Crown, 1988; Elias, Hayes, & Broerse, 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, Bobb, & Clark, 2016; Hilbrink et al., 2015; Lourenço, Pereira, Sampaio, & Coutinho, 2021; Stivers, Sidnell, & Bergen, 2018).

Hilbrink et al. (2015) studied the frequency and duration of gaps and overlaps in free-play interactions with 12 mother-infant dyads, from Wales, at 3, 4, 5, 9, 12, and 18 months. Infants produced above one third of their turns in overlap, but by 18 months their proportion was closer to their mothers', at 20%. Infants' median gap durations ranged around 600 ms from 3 to 5 months, increased at 9 months to around 1100 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 duration from 3 to 5 months, has very little change, slightly varying around 500–600 ms, and finally 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 from Portugal. The effect of face-to-face (free-play without toys) and object-oriented interactions (free-play with toys; challenging-toy play) in turn-transition duration was also considered. Infants' median gap durations only slightly increased over time for the object oriented tasks, from 420 ms *($M = 1012$ ms)* to 570 ms *($M = 1106$ ms)* in the free-play with toys task, and 430 ms to 490 ms in the challenging-toy play task. Inversely, median gap durations slightly decreased on the free-play without toys task, from 301 ms *($M = 630$ 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 of turn-transition interaction, were found for gap duration. Infant gaps – gaps in mother-to-infant turn-transitions, i.e. where the infant is the partner that gains the floor – were significantly longer than their mothers. When toys were removed from the interaction, gaps in mother-to-infant turn-transitions were shorter than in object-oriented tasks, but also followed the same descending trajectory as in their mothers, with shorter gaps at 12 months compared to 7 months. In contrast, overlap duration in mother-to-infant turn-transitions 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.

Lourenço, Pereira, et al. (2021) 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 throughout that period, and the possible interference of the object manipulation and exploration 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 later in development, when linguistic processing becomes more complex (Lourenço, Pereira, et al., 2021).

Casillas et al. (2016) analyzed the response latency in turn transitions that were also pair-adjacent (question-responses pairs). Data was selected from the naturalistic conversations of 5 North-American 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. 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 ($Mdn = 651$ ms) 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). This 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, Gambi, & Rabagliati, 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 from Canada. 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, Versyp, Cox, & Fusaroli, 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.

1.2. 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. Interactive contingency in turn-taking is often compared to a glue holding turns together, and constitutes the coordinative basis for effective communication (Casillas, 2014). When distinguishing between temporal and semantic contingency, the latter builds upon the temporal relations of the former, by framing them onto semantically meaningful verbal transactions (Casillas, 2014). From a microanalytic perspective though, interactive contingency is fundamentally related to predictability and can be reliably measured from the temporal relations of turn-transition timing, regardless of the level of linguistic development (Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001). Analysis of the whole interaction provides a measure of temporal contingency that is unconstrained by a specific semantic relationship, and therefore be comparatively applied to pre-verbal and verbal children.

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 value (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 these 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 min free-play task, where toys were available and conversation was permitted.

We measured all vocal exchanges of the dyads (the final dataset consisted of $N = 7113$ turn-transitions) 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 better estimate of the distribution of children's turn-transition timing, one that is not constrained by a specific semantic relationship, and can therefore be used as reference for comparing children's and infants' turn-transition timing.

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.

2. Method

2.1. 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.

2.2. Procedure and materials

Interactions unfolded 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, within arm reach of the participants, and 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, connected 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 min. 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.

2.3. Coding of gaps and overlaps in turn-transitions

For the coding of mother and infant turn-transitions, each vocalization (made by the mother or the child) was segmented. First, the onset and offset of each vocalization event was automatically generated using Praat (Boersma & Weenink, 2022), and this information exported. Second, the individual vocalization event timestamps (onset, offset) were imported into ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Third, the onset and offset of each individual vocalization event was verified, and manually adjusted using ELAN, if necessary, by research assistants, to ensure the accuracy of the segmentation.

Assistants were trained to segment vocalizations, in two phases. They first received instruction on how to use ELAN and manually identify the start and stop of vocalization events in training videos, supervised by an experienced annotator. In a second step, the assistants were instructed on how to adjust the onset and offset of the segmentations generated automatically by Praat, according to their previous training. The resulting segments were then automatically coded using custom developed Matlab scripts that followed the on-off binary logic of the Automated Vocal Transaction Analyzer (AVTA; Cassotta, Feldstein, & Jaffe, 1964). The input to the AVTA model are two binary time series, each representing the vocalization events of the dyad members. This model makes use of a turn-transition rule and a set of dyadic states (i.e. the state is jointly determined by the dyad members). The turn-transition rule states that there is a turn holder, and that a turn-transition occurs when the turn holder goes silent and the other partner vocalizes alone. This produces five possible states: (1) the partner that is the turn holder either vocalizes (Vocalize state) or (2) is silent (Pause state); during a turn-transition either there is a silence, i.e. a gap, and this is (3) the switching pause (SP state) or there is (4) an overlap (an interruptive simultaneous speech, the ISS state); finally, when the partner that is not the turn holder vocalizes but goes silent with a turn-transition occurring, this is (5) a non-interruptive simultaneous speech (NSS). For an in-depth presentation of the AVTA model please the methods section of Jaffe et al. (2001) monograph and Lourenço, Pereira, et al. (2021) where the same method is used. The resulting dyadic states of interruptive simultaneous speech (ISS) and switching pause (SP) correspond, respectively, to overlaps and gaps.

2.4. 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) that regains the floor. For convenience we will refer to partner (mother or child) in the sense of what partner regains the floor, so mother refers to child-to-mother turn-transitions and child refers to mother-to-child turn-transitions. We also calculated average and median floor-transfer offset (FTO), gaps and overlaps durations, by partner for the three age groups. Both mean and median durations are reported, as such facilitates comparisons between different studies. As we have modeled mean durations, we primarily report mean durations, with median durations reported in parentheses.

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 test were kept.

Models were fitted using the *lmer* and *glmer* functions of the *lme4* package (Bates, Maechler, & Dai, 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).[1]

3. Results

3.1. Descriptive statistics

3.1.1. Frequency

By applying the AVTA model to the segmented vocalizations a total of 7208 turn-transitions were detected, of which 1.3% were excluded from the analysis, for being 3.5SD longer (4954 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. This is a conservative cut-point; for comparison, Casillas et al. (2016) used 2.5 SD. From the remaining 7113 turn-transitions, 3570 were produced by the children, and 3453 were produced by their mothers. Table 1 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 compared with those expected in adults (Heldner & Edlund, 2010).

3.1.2. 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.

Fig. 1 shows the distribution of gaps and overlaps duration.

Table 1
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%

Note. For convenience, the partner that regains the floor is listed in the column labelled Partner, so Child means mother-to-child turn-transitions.

[1] The dataset and data analysis code is available at https://osf.io/82rbd/.

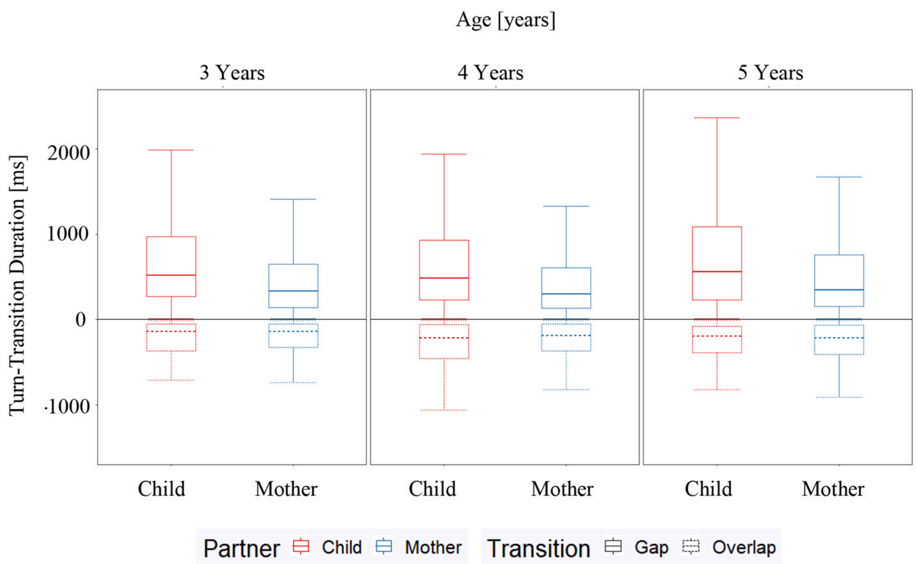

Fig. 1. Boxplot of the duration [ms] of gaps and overlaps duration by partner and age group.
Note. For convenience, we label the member of the dyad that regains the floor, during a turn transition, as Partner; e.g. Child means mother-to-child turn transitions. Transition is either a gap or an overlap where Gap means a silence during the turn-transition and Overlap means that onset of the vocalization of who regained the floor overlapped with the end of the other partner's vocalization.

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 distribution of turn-transition duration, measured as FTO, illustrated in Fig. 2, had a slightly skewed and unimodal distribution, similar to what is expected in adults (Levinson & Torreira, 2015; Stivers et al., 2009). Density functions estimations, show that the distribution of turn-transition approximates to a Normal distribution. We also inspected the Q-Q plots and found that distribution was approximately Normal, with minor violations at the tails. The same is not to be 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.

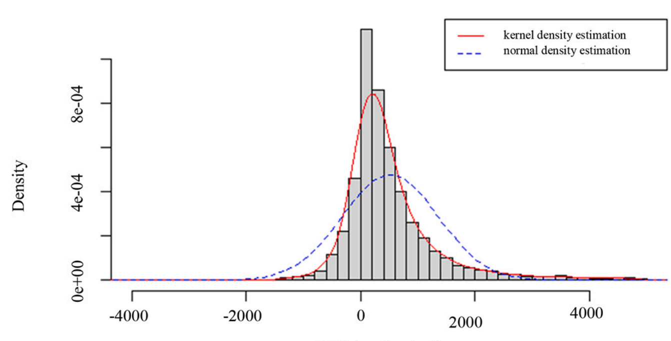

Fig. 2. Histogram of turn-transition duration (measured as FTO).
Note. A kernel and normal density estimation, using default parameters, is overlaid.

3.2. *Modeling of turn-transition duration*

In line with the delineated research objectives, turn-transition duration was modeled first as floor-transfer offset (FTO). After, gaps and overlaps were considered separately.

3.2.1. *Floor-transfer offset (FTO)*

To understand the predictor effect of child's age and direction of turn-transition (child, i.e. mother-to-child turn-transition; mother, i.e. child-to-mother turn-transition) 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 Fig. 2), 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, $\chi^2(1) = 142.63$, $p \leq 0.001$. 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 2 shows the results for the final model (log-likelihood = −8631.2, $N = 7113$).

There was no significant main effect of age in dyadic FTO duration ($\beta = 1.08$, CI = 0.98–1.19, $p = 0.101$), with Fig. 3 showing only a slight increase in FTO duration throughout age points.

On the other hand, a significant main effect was found for the direction of turn-transition ($\beta = 0.12$, $CI = 0.10–0.13$, $p < 0.001$). Fig. 4 plots the estimated marginal means and confidence intervals for the predictor effect of direction, showing that mother-to-child turn-

Table 2
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

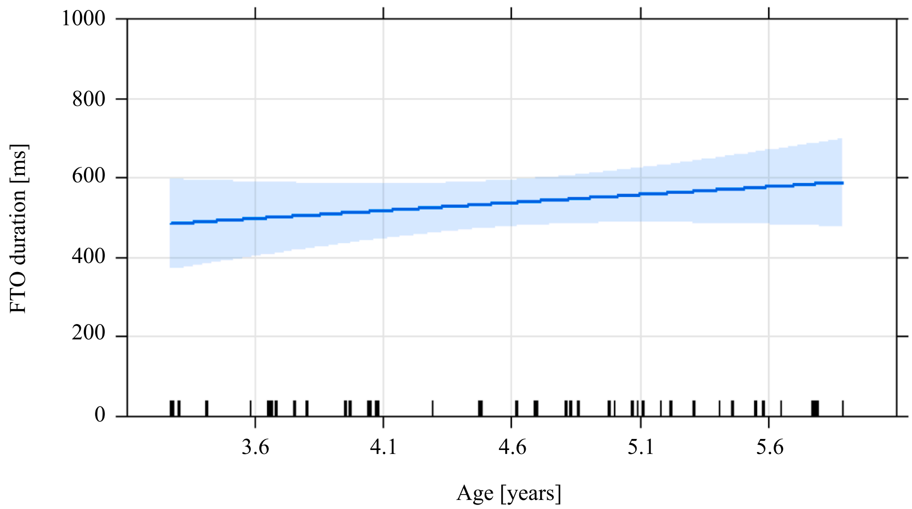

Fig. 3. Predictor effect plot for age in the FTO duration model.

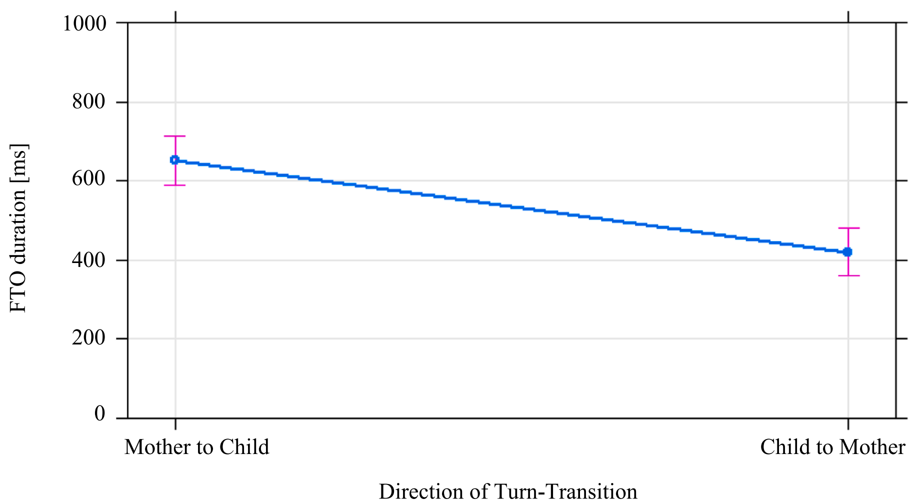

Fig. 4. Predictor effect plot for direction in the FTO duration model.

transitions are, on average, longer (M = 650 ms) than the gaps in the opposite direction, i.e. child-to-mother (M = 420 ms).

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.

3.2.2. 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 distribution of gaps and overlaps.

We first built a base model with gap duration as a dependent variable and a random intercept per dyad, as a random effect. We then compared it 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, $\chi^2(1) = 125.76$, $p < 0.001$. 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 3 shows the results for the final model (log-likelihood = −3529.1, N = 5831).

Table 3
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 ($\beta = 1.08$, $CI = 0.98$–1.19, $p = 0.101$) was found. Nevertheless, Fig. 5 still shows a more accentuated increase tendency over time in dyadic gap duration, than with FTO duration.

Similarly to FTO, a significant main effect was found in gap duration for the direction of turn-transition ($\beta = 1.16$, $CI = 1.13$–1.19, $p < 0.001$). Fig. 6 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).

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

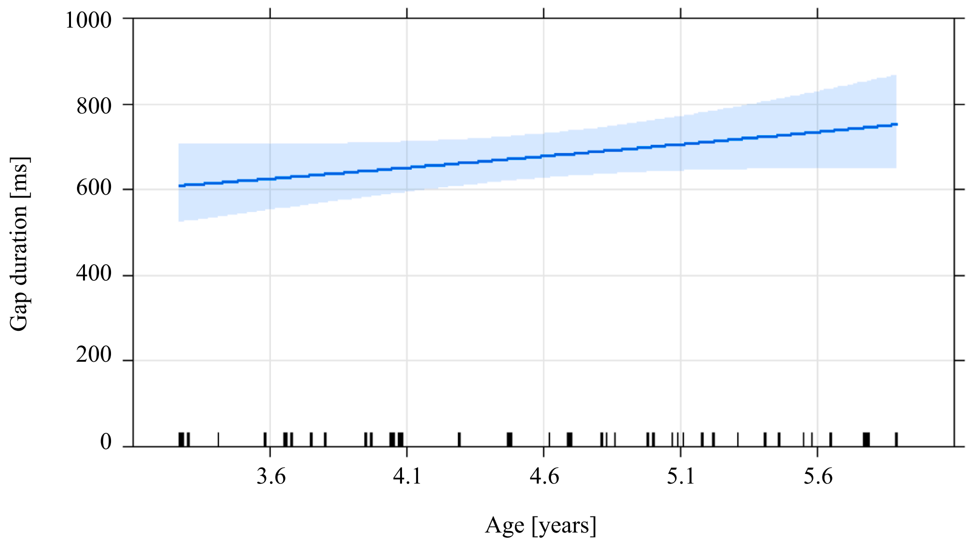

Fig. 5. Predictor effect plot for age in gap duration.

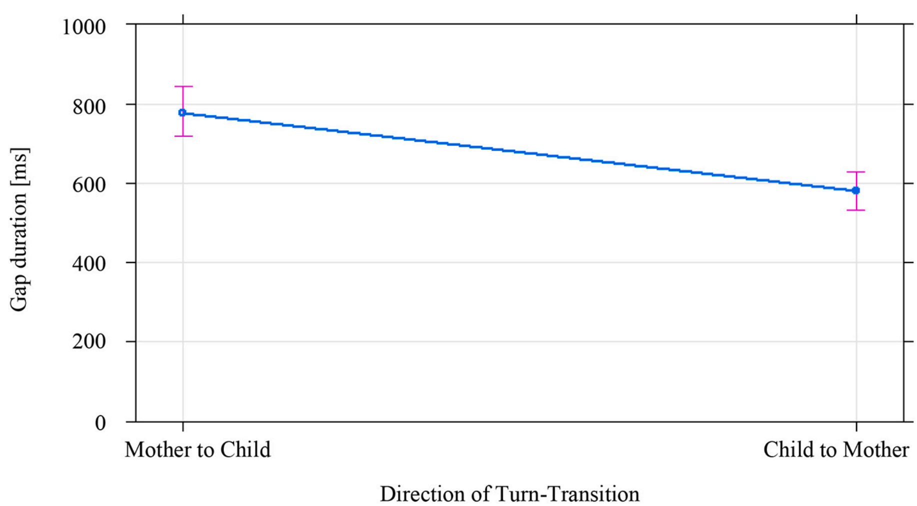

Fig. 6. Predictor effect plot for direction in the gap duration model.

turn-transitions independently. Again, none of the explorations provided evidence of any significant effect of age in gap duration.

3.2.3. 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 4 shows the results for the final model (log-likelihood = 330.60, $N = 1282$).

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 Fig. 7 that shows a much more flattened evolution from age point to age point than with gaps, or even FTO durations.

Table 4
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

Likewise, there was no significant main effect of the direction of turn-transition in overlap duration ($\beta = 1.04$, $CI = 1.13$–1.19, $p = 0.229$), as illustrated by Fig. 8.

4. 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 of turn-transition

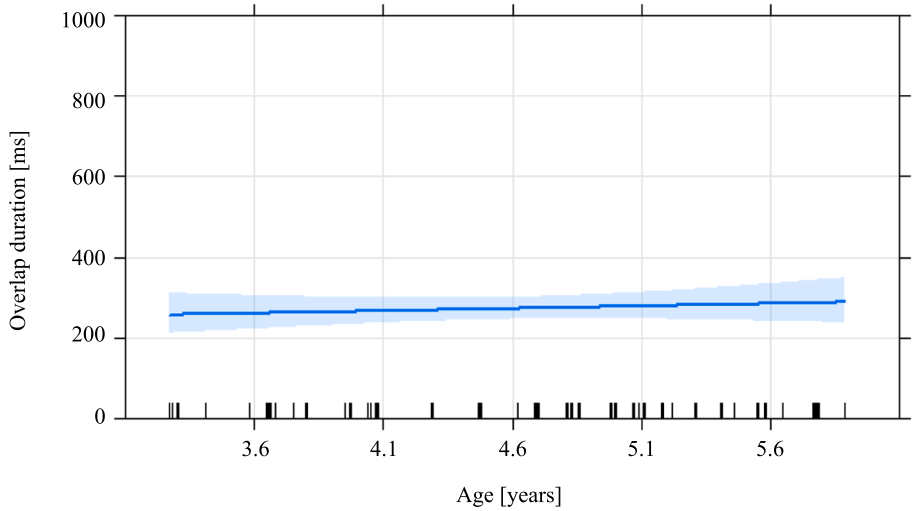

Fig. 7. Predictor effect plot for age in the overlap duration model.

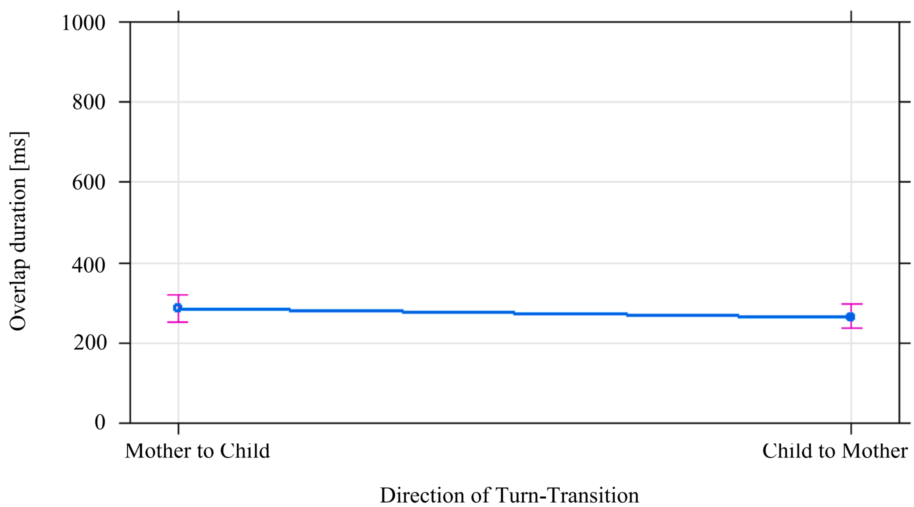

Fig. 8. Predictor effect plot for direction in the overlap duration model.

timing in 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.

4.1. 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 when measuring turn-transition duration through the FTO, durations of the whole interaction are still comparable to the timings reported by research that examined only semantically contingent turn-transitions.

Direct comparison of our findings with findings from studies of turn-taking in mother-infant interactions is difficult because of methodological and age differences. Nevertheless, 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.

4.2. 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.

5. General discussion and 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.

In terms of the general theoretical implications of the slow pattern of progression of turn-transition timing (Casillas et al., 2016; Stivers et al., 2018), we interpret it as an indication of the stability of the turn-taking system. For one, overlap timing appears to be already stabilized throughout this period. This bears out not only through our analysis that found no age differences in overlap duration as well as no significant differences between mother-to-child and child-to-mother overlap timings, but also when comparing mother-to-child timing with the average overlap duration (275 ms) of adults (Heldner & Edlund, 2010). The stability of the system can also be noticed from the timing of gaps. Notwithstanding the diversity between the linguistic repertoire of 3, 4 and 5-years-olds, the turn-taking system appears also to have capped around 700–800 ms throughout this period. This suggests that even if more complex linguistic comprehension and production is being employed by older children, the system may be integrating it by converging towards the established timing of the system. It is also noteworthy to point out that the adult-adult timing, that we are using here as a reference, has noticeable differences with the child but also with the mother, in the sense that child-to-mother turn-transition duration is also slower than in adult-adult conversational turn-taking. A question for future studies should therefore be how to disentangle the role of the mother, in terms of sensitively adjusting to the child's preferred rhythms, from the child's full competencies.

By measuring gaps and overlaps separately we have shown that there are still differences in gap and overlap duration consistent with the work by Hilbrink et al. (2015) and more recently Lourenço, Pereira, et al. (2021). We have also suggested that considering the timings reported in those studies and our results, overlaps may still be getting shorter, before the 3-years-old mark, towards the standard overlap timing of adults. We have suggested that these findings provide confirmatory evidence that adult-like overlaps may indeed be acquired not only developmentally earlier than adult-like gap durations, but considerably earlier.

There are several reasons for that to be the case. If we follow Levinson and Torreira (2015) discussion of the cognitive implications of the standard model of turn-taking (Sacks et al., 1974), turn-transition timing challenges even the production timing of the shortest turn, which implies that comprehension and production processes are overlapping to an extent to produce such short-turns. To produce short gap durations requires already such a cognitively challenging processing, that for overlaps – if conceived as inverse gaps – to occur it requires an even greater amount of superposition between comprehension and production processes, that can only be accounted for through predictive processing (Levinson & Torreira, 2015). Indeed, overlaps still occur in adults, although they are less frequent that gaps and usually, no longer than 200 ms (Heldner & Edlund, 2010).

More importantly, the target of turn-taking timing is not only to minimize gaps, but also avoiding overlaps (Sacks et al., 1974). This bears out evolutionarily, as the communication processes of the human species would be severely affected if the interlocutors would attempt to comprehend the other's message and produce their own's simultaneously (Levinson, 2006; Sidnell, 2001). This points towards the sequential organization of turn-taking that has also been demonstrated in other primate and non-primate species (Levinson, 2019). We suggest that in this sense, overlaps not only present strong evidence of the superposition between comprehension and production processes, but also demonstrate, by simply occurring when the entire system of turn-taking is structured around minimizing them, how challenging it still is to accurately predict when the present turn will end, and when to launch the next turn. This predictive ability is an essential component of the Interaction Engine Hypothesis (Levinson, 2006, 2019) and assumed to be developed early in development. We have discussed how observational and experimental studies have shown that pre-verbal infants pairtake as early as 1.5 months in the sequential organization of turn-taking (Bateson, 1975; Jasnow & Feldstein, 1986), present a sensitivity to disruptions of interactive contingency at least by 3 months (Murray & Trevarthen, 1985; Striano et al., 2006), and can accurately predict turn ending (Casillas & Frank, 2013) and distinguish between communicative intentions already in their first year (Thorgrimsson, 2014). We have also seen in Hilbrink et al. (2015) and in Lourenço, Pereira, et al. (2021), as well as in our results that children and infant produce overlaps around the same frequency that adults.

Therefore, we suggest that the coordination and predictive abilities required to minimize overlaps and overlap durations develop earlier and are consolidated by the 3-years-old mark. While the ability to minimize gap durations is more dependent on a more harmonious integration between the interaction engine system and the linguistic system, that is not yet attainable at all levels of linguistic complexity by 5-years-olds.

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 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.

CRediT authorship contribution statement

Vladimiro Lourenço: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Visualization, Funding acquisition, Writing – original draft, Writing – review & editing. **Juliana Serra:** Methodology, Software, Validation, Investigation, Writing – review & editing, Visualization, Data curation. **Joana Coutinho:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Alfredo F. Pereira:** Conceptualization, Methodology, Software, Formal analysis, Resources, Data curation, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Data availability

The dataset and data analysis code is available at https://osf.io/82rbd/.

Acknowledgements

This study was conducted at the Psychology Research Centre (CIPSI/01662), School of Psychology, University of Minho, supported by the Foundation for Science and Technology (FCT) through the Portuguese State Budget (UIDB/PSI/01662/2020). This research was also supported by FCT doctoral fellowship PD/BD/128465/2017 awarded to Vladimiro Lourenço, FCT projects IF/00217/2013, PTDC/MHC/PCN/1530/2014, PTDC/PSI-GER/2463/2021 awarded to Alfredo F. Pereira, and Bial Foundation project 87/12 awarded to Joana Coutinho. We sincerely thank the generosity of the following institutions in Guimarães that supported the project (Cisdec, Fraterna, Patronato Nossa Senhora da Oliveira, Centro Social de Azurém, Venerável Ordem de São Francisco, Infantário Nuno Simões, and Centro Social da Paróquia de Penselo). Finally, we acknowledge our participants (parents and children) without whom this study would not have been possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105568.

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

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer. [Computer program] version 6.3.03. retrieved 17 December 2022 from http://www.praat.org/.

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. D. (2015). Mother-infant contingent vocalizations in 11 countries. *Psychological Science, 26*(8), 1272–1284. https://doi.org/10.1177/0956797615586796

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. In *Proceedings of the 35th annual meeting of the cognitive science society* (pp. 299–304). http://langcog.stanford.edu/papers/CF-cogsci2013.pdf.

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

Elias, G., Hayes, A., & Broerse, J. (1986). Maternal control of co-vocalization and inter-speaker silences in mother-infant 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. In *Studies in mother-infant interaction, January 1977* (pp. 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

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), 1–132. i–viii http://www.ncbi.nlm.nih.gov/pubmed/11428150. i–viii.

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), Article 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*, 1–16. https://doi.org/10.1037/pne0000276. September.

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 Computing. 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.

Sidnell, J. (2001). Conversational turn-taking in a Caribbean English creole. *Journal of Pragmatics, 33*(8), 1263–1290. https://doi.org/10.1016/S0378-2166(00)00062-X

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., … 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. In *, 7(2). Interaction studies. Social behaviour and communication in biological and artificial systemsinteraction studies / social behaviour and communication in biological and artificial systems interaction studies* (pp. 251–271). https://doi.org/10.1075/is.7.2.08str

Thorgrimsson, G. B. (2014). *Infants' understanding of communication as participants and observers. [doctoral thesis, Radbound university Nijmegen]*. Radbound Repository. https://repository.ubn.ru.nl/bitstream/handle/2066/133676/133676.pdf.

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. In *Proceedings of the 5th international conference on language resources and evaluation, LREC 2006, August* (pp. 1556–1559).