DOI: 10.1111/cdev.13754

CHILD DEVELOPMENT

REVIEW

# A systematic review and Bayesian meta-analysis of the development of turn taking in adult–child vocal interactions

**Vivian Nguyen**[1,2] | **Otto Versyp**[1,2] | **Christopher Cox**[2,3,4] | **Riccardo Fusaroli**[2,3,5]

[1]Psychology, Ghent University, Gent, Belgium

[2]Department of Linguistic, Cognitive Science and Semiotics, Aarhus University, Aarhus, Denmark

[3]The Interacting Minds Center, Aarhus University, Aarhus, Denmark

[4]Department of Language & Linguistic Science, University of York, York, UK

[5]Linguistic Data Consortium, University of Pennsylvania, Philadelphia, Pennsylvania, USA

**Correspondence**
Riccardo Fusaroli, The Interacting Minds Center, Aarhus University, Aarhus, Denmark.
Email: fusaroli@gmail.com

**Funding information**
National Institute on Deafness and Other Communication Disorders, Grant/Award Number: R01; Interacting Minds Center

**Abstract**

Fluent conversation requires temporal organization between conversational exchanges. By performing a systematic review and Bayesian multi-level meta-analysis, we map the trajectory of infants' turn-taking abilities over the course of early development (0 to 70 months). We synthesize the evidence from 26 studies (78 estimates from 429 unique infants, of which at least 152 are female) reporting response latencies in infant–adult dyadic interactions. The data were collected between 1975 and 2019, exclusively in North America and Europe. Infants took on average circa 1 s to respond, and the evidence of changes in response over time was inconclusive. Infants' response latencies are related to those of their adult conversational partners: an increase of 1 s in adult response latency (e.g., 400 to 1400 ms) would be related to an increase of over 1 s in infant response latency (from 600 to 1857 ms). These results highlight the dynamic reciprocity involved in the temporal organization of turn-taking. Based on these results, we provide recommendations for future avenues of enquiry: studies should analyze how turn-by-turn exchanges develop on a longitudinal timescale, with rich assessment of infants' linguistic and social development.

## INTRODUCTION

Conversations involve structured and coordinated interactions between two or more interlocutors who *take turns* in their roles as speakers and listeners (Fusaroli, Rączaszek-Leonardi, et al., 2014). The fluency of these conversational exchanges depends on their dynamic temporal structure where each speaker has to anticipate the end of the other speaker's turn (Sacks et al., 1978). To keep the conversation running smoothly, interlocutors tend to optimize the response latency between their turns by avoiding overlaps in speech and minimizing the pauses between their own utterances and those of their partner. Turn-taking is schematically represented in Figure 1. Note that this turn-taking system represents a coarse-grained and yet useful description of conversations. At a higher level of detail, conversations involve frequent backchanneling (nodding and supporting the speaker with sounds like "hmm") and other subtle ways of communicating beyond a strict speaker/listener distinction (Cummins, 2013).

This tendency for interlocutors to optimize the latency between their utterances applies in a wide variety of human languages. Stivers et al. (2009), for example, investigate ten typologically distinct languages and show that—in adult–adult conversations—average response

**Abbreviations:** CI, compatibility interval; ER, evidence ratio; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Vivian Nguyen, Otto Versyp, and Christopher Cox joint first-authorship.

© 2022 The Authors. *Child Development* © 2022 Society for Research in Child Development.



14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

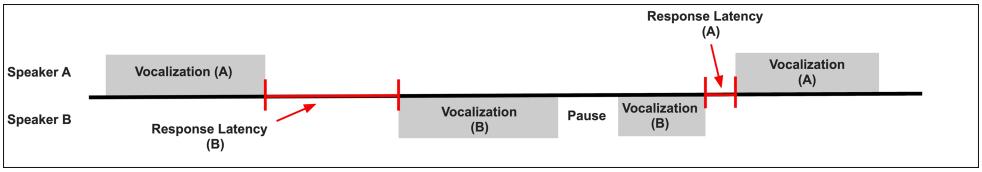


**FIGURE 1** Visualization of the turn-taking process. The *x*-axis represents time. Pause refers to within-speaker pauses. Response latency—the focus of the current study—is the period of time which the corresponding speaker requires to answer an utterance by another speaker (i.e., duration of the turn transition to another speaker). Note that response latency could be negative, if the adjacent turns overlap

latency during turn transitions varies between 0 and 500 ms. The cross-linguistic variation in this corpus is non-trivial, with latencies averaging 7 ms in Japanese against the 470ms of Danish. However, the range of acceptable latencies is often overlapping, and conversational turn-taking exhibits similar qualitative structures across languages. Notably, these latencies are most often too short to allow for speech planning to happen after one's interlocutor has finished speaking (Griffin & Bock, 2000; Indefrey, 2011; Indefrey & Levelt, 2004). This indicates that speakers are well-tuned to transitions between turns, to the point that they can anticipate them and start planning before the end of their interlocutor's utterance. Turn-taking has thus been argued to provide the cognitive scaffolding to facilitate social interactions (Stivers et al., 2009).

Furthermore, a growing number of studies is finding signs of communicative turn-taking within different species, often sharing the optimization principles highlighted in human conversation (Pika et al., 2018). For example, nightingales are capable of precisely timing the onset of their song, starting maximally 1 s after the end of a neighbor's song (Todt & Hultsch, 1999). Other studies have identified vocal turn-taking in meerkats (Demartsev et al., 2018) and some primate species (e.g., *Lepilemur edwardsi*, Méndez-Cárdenas & Zimmermann, 2009). These findings have been used to suggest that turn-taking is one of the basic foundations of social communication across a variety of different species (de Reus et al., 2021).

Perhaps because of its foundational basis for social interaction, turn-taking plays an important functional role in human language development. For instance, pre-verbal vocalizations more closely resembling language have been shown to elicit prompter caregiver responses; and in turn, faster caregiver response latencies are shown to relate to more language-like successive vocalizations (the "social feedback loop" model, Lopez et al., 2020; Warlaumont et al., 2010, 2014). Furthermore, at least in adult–adult conversations, smooth and regular turn-taking may play a role in facilitating coordination. For instance, synchrony and shared tempo in joint activities have been shown to create feelings of shared motivation and to lessen the cognitive demands required by the interaction, since it reduces the need to continuously adjust to changing conversational rhythms and to figure out when to talk (Cross et al., 2020; Dideriksen et al., 2020; Fusaroli, Konvalinka et al., 2014; Fusaroli, Rączaszek-Leonardi, et al., 2014; Fusaroli & Tylén, 2016; Fusaroli et al., 2016; Konvalinka et al., 2011).

While this turn-taking process appears to be effortless, it relies on a set of complex skills (e.g., simultaneous language processing and production planning) and coordination between the different interlocutors. The order in which speakers organize their conversations, the length of their turns, and the pauses between those turns all require real-time coordination and cannot be defined in advance. Furthermore, interlocutors need to process the other person's utterance, plan their own response in accordance while simultaneously anticipating when the end of a turn is coming up (Pickering & Garrod, 2021; Riest et al., 2015; Ruiter et al., 2006).

Despite the skills required, infants seem capable of smooth turn-taking at an early point in development (Bateson, 1975; Gratier et al., 2015). However, there is no current overview on how early this might happen and more generally of the developmental trajectory of turn-taking. Reports of response time latencies in parent–child interactions are scattered through the literature, often reported as secondary descriptions of the data. Understanding how infants' turn-taking abilities develop over time would grant us a more complete understanding of language development and the skills underlying it, as well as provide insights as to the cognitive mechanisms necessary for turn-taking.

To provide better foundations for investigations of turn-taking, this paper systematically reviews and meta-analyzes the evidence related to the development of infants' turn-taking abilities. This paper will specifically focus on infant–adult dyadic interactions (i.e., no infant–infant interactions), since adult interlocutors provide more homogeneously skilled interlocutors than other infants and therefore might facilitate comparisons across infant ages.

In the following sections, we present models of turn-taking and thereafter summarize the literature on how turn-taking develops over the course of early infant development. We then discuss the factors that likely moderate turn-taking in order to motivate the specific hypotheses to be assessed in the meta-analysis.

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

## Models of turn-taking

Turn-taking is construed as being organized around two key principles: (i) to ensure minimal overlap between speakers' utterances, and (ii) to optimize the transition times between speakers' conversational turns (Stivers et al., 2009). Several mechanisms have been suggested to make this possible.

In one of the earliest in-depth analyses of conversational turn-taking, Sacks et al. (1978) developed a descriptive model of turn allocation, that is, of the selection of the next speaker. As an utterance ends, three things can happen: (i) the current speaker selects who should speak next; (ii) a current listener self-selects and starts speaking; or (iii) the current speaker self-selects and resumes speaking. Independently of which of these options is taken, Stivers et al. (2009) have argued for smooth transitions to happen universally across humans by showing modal response time between 0 and 500 ms across 10 distinct languages. This is argued to indicate a universal mechanism aimed at both avoiding speech overlap and minimizing periods of silence between turns (Levinson & Torreira, 2015). While the general pattern is uncontroversial, it has been argued that a modal range of 0–500 ms also indicates potential non-trivial cross-linguistic differences. For instance, Danish presents response latencies that are 100 ms slower than average, and acceptable response latencies up to 1s, besides other differences (Heinemann, 2010). This has been associated with compensation mechanisms: Danish presents unusually high levels of phonetic reduction making the speech signal potentially more difficult to parse than many other languages (Trecca et al., 2021). However, these cross-linguistic differences have not been systematically investigated further.

Levinson (2006, 2016) takes an explicitly cognitive direction to explain smooth turn-taking, with the Interaction Engine Hypothesis. This model focuses on the abilities necessary for turn-taking, stressing that such interactional abilities are also foundational for language development and use (Levinson, 1995; Warlaumont et al., 2014). These cognitive abilities include the crucial components of the engine: attributing intentions to others (mapping their behavior onto inferred goals); simulating an interlocutor attributing intentions to us (anticipating what an interlocutor thinks we think; Waade et al., 2020); and producing behaviors aimed at having interlocutors recognize the speaker's intentions (or Gricean intentions). Crucially, this model provides specific predictions, namely that smooth fast turn-taking would be present in pre-verbal infants, but that it would slow down with the acquisition of language (which the author puts at around 12 months of age), as developing linguistic skills would need to be integrated to the existing interactional ones. One study has tested and partially supported this prediction, highlighting a slow-down in turn-taking at around 9 months, that is, before the actual use of words, (Hilbrink et al., 2015).

If the Interaction Engine Hypothesis focuses on the individual cognitive abilities needed to make interactions work, Wilson and Wilson (2005) focus more explicitly on the interpersonal dynamics of turn-taking, with the Coupled Oscillators model. The model states that interlocutors have internal oscillators (e.g., rhythmic patterns of brain activity), which can become progressively coupled across individuals through coordination of response latencies and syllable durations. Turn-taking can be smooth because it is based on the rhythm of these oscillators. For instance, we can imagine interlocutors establishing a shared tempo of 500 ms per oscillation, so that response latencies would only happen at multiples of 500ms and display the same durations: 0 s, 500 ms, 1 s, etc. This not only simplifies the complexity of when to start speaking (as only specific timings are preferred), it crucially minimizes the odds of speaking simultaneously due to the fact that the readiness cycles of listeners are counter-phased to those of speakers. This model is supported in a few animal turn-taking studies (Takahashi et al., 2013). However, evidence for the model in human conversations is controversial. In a case study of one conversational pair, O'Dell et al. (2012) failed to detect relations between syllable rate and turn-taking and suggested that slower speech-related rhythms, such as phrasal stress rhythm, might be involved. Alternatively, McFarland (2001) and Rochet-Capellan and Fuchs (2014) explore the possible role of respiration and indicate the presence of at least two respiration regimes that get interpersonally entrained: (i) within turns, and (ii) during turn transitions. Despite the uncertainty as to the specific details of the mechanisms, growing evidence shows that interlocutors engaged in conversation adjust the temporal aspects of their speech to each other, for example, the duration of intra-turn pauses (Cappella, 1981); the duration of speech and pause sequences (Fusaroli & Tylén, 2016; Fusaroli et al., 2016); and the tempo of their speech and syllables (Manson et al., 2013); as well as the temporal aspects of non-speech behaviors (Curioni et al., 2019; Duran & Fusaroli, 2017; Fusaroli et al., 2016; Pezzulo et al., 2019).

While there is currently no consensus on the specific mechanisms underlying turn-taking, we should expect cognitive development, speech development, and interpersonal dynamics to play a role in both the developmental trajectory and deployment of turn-taking. We now turn our gaze to current evidence as to how turn-taking develops from early infancy.

## The developmental trajectory of turn-taking and its moderators

The earliest described examples of vocal turn-taking in infants are called proto-conversations. Proto-conversation

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

can be described as a structured, nonlinguistic vocal exchange between infants and their peers or adults (Hsu & Fogel, 2003; Hsu et al., 2001; Kaye, 1977), reported already before 3 months of age (Bateson, 1975; Dominguez et al., 2016; Gratier et al., 2015). These proto-conversations are described as shared meaningful activities, and seem to adhere to the temporal structure of adult conversations (Bateson, 1975; Snow, 1977). While proto-conversations have been characterized as analogous in response latencies to adult conversations, older infants already from 9 months of age, and up to 4 years of age have shown to be slower than adults at planning, initiating, and completing their responses (Casillas et al., 2016; Hilbrink et al., 2015; Levinson, 2013).

The mechanisms underlying the ability to sustain turn-taking and thus the reported patterns are still under-investigated. There is evidence that infants are sensitive to turn-taking timing in social interactions at least from 9 months of age (Striano et al., 2006; Thorgrimsson et al., 2014), and that at least from the age of 1 year infants are capable of predicting upcoming turn transitions (Casillas & Frank, 2013). More direct tests that manipulate the degree of predictability of questions directed to older (3;0–6;0) children reinforce this picture: the more predictable the question, the more children shorten their response latency (Lindsay et al., 2019). This finding admits an important role of the contexts and activities at play in scaffolding children's turn-taking abilities.

However, very few longitudinal studies of turn-taking have been conducted, and none investigating the joint development of response latencies and their underlying mechanisms (Bateson, 1975; Ginsburg & Kilbourne, 1988; Hilbrink et al., 2015; Snow, 1977).

## Potential factors underlying turn-taking skills

In order to develop a more comprehensive understanding of the development of turn-taking, we reviewed the literature to identify possible factors to be accounted for. This review allowed us to investigate the following moderators in the meta-analysis: (i) child age, (ii) adult response latencies, (iii) familiarity of interlocutors and (iv) interactional settings, (v) type of activity, (vi) language being spoken, and (vii) whether the child had a developmental disorder or other special conditions. We discuss each in turn.

### Age

The Interaction Engine Hypothesis (Levinson, 2006) predicts turn-taking to be fast at an early age, but to slow down after 12 months. Tomasello (1999) on the contrary talks about a "9-month revolution", the age at which infants start recognizing the importance of intentional communication, which might also cause a slowdown (Hilbrink et al., 2015). This slowdown can additionally be attributed to the fact that infants at this age are rapidly acquiring more complex language, which they attempt to incorporate in their turns. The empirical results from longitudinal studies are, however, less straightforward. For example, testing the Interaction Engine Hypothesis, Hilbrink et al. (2015) found that up until 9 months preverbal infants' response latencies are similar to those of adults, but slow down thereafter, until infants are 18 months of age. Thus, the authors amend the original prediction of a slowdown from 12 months, and place it at 9 month, referring to Tomasello's 9-month revolution (Tomasello, 1999). Another study by Casillas et al. (2016) reported that slowdowns and speed-ups might be relative to the specific context: response latency speeds up earlier for answering simpler questions. However, a cross-sectional study by Gratier et al. (2015) reported a slowdown already from 2 to 3 months to 4- to 5-month-old infants. All these findings seem to suggest that age might play a role in turn-taking dynamics, but also that the developmental trajectory of turn-taking is complex and not yet adequately investigated. Based on the aforementioned literature showing that infant response latencies slow down before speeding up again, we expected to find a quadratic relation between the age of infants and their response time latencies.

### Adult response latency

As we have argued, turn-taking is an interpersonal activity, where individuals adjust to each other in an interactive fashion (Duran & Fusaroli, 2017; Fusaroli, Rączaszek-Leonardi, et al., 2014; Fusaroli et al., 2021). Infants and children would then be likely to adjust to adult response latencies (and vice versa). Accordingly, several studies present evidence for 4- and 9-month-old as well as 4- to 5-year-old infants adjusting their response latencies to those of adults and vice versa (Beebe et al., 1988; Jasnow & Feldstein, 1986; Welkowitz et al., 1990). Furthermore, direct experimental manipulations showed that when parents delayed their answers by pre-agreed latencies (1 and 3 s), their 4-year-old children produced corresponding latencies (Newman & Smit, 1989). These studies lead us to expect an effect of adult response latencies: slower adult response latency will correspond to slower infant response latency and vice versa.

### Familiarity of interlocutors and settings

Children have been shown to adjust their response latencies to the predictability of the utterances they were responding to. One could hypothesize that the predictability of the situation (and therefore its familiarity) would also play a role. Indeed, infants are reported to vocalize in a more contingent way when interacting with

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a familiar partner than with an unfamiliar one (Bigelow, 1998). However, 4-month-old infants have been shown to display shorter response time latencies when they are interacting with strangers than with familiar interlocutors (Jaffe et al., 2001). This latter result would fit with findings in adults, reporting that interactions with unfamiliar interlocutors result in increased vocal activity and more tightly knit turn-taking coordination compared to familiar ones, and that this is even more salient when one of the adults has aphasia (Cappella, 1996; Crown, 1991; Dressler et al., 2009). A possible explanation for the inconsistent findings may include that infants are not reacting to the familiarity of the interlocutor, but to more specific adult behaviors. Indeed, children adjust to the response latencies of slower and faster unfamiliar adults (Bigelow, 1998). The effects of familiarity of other components of the conversational settings—for example, familiar (i.e., home, daycare) versus unfamiliar settings (i.e., laboratory, clinic)—have not been directly tested. However, vocal behaviors have been observed to vary: 4- to 5-month-old children produce double the amount of vocalizations when at home compared to laboratory settings (Lewedag et al., 1994), while mothers vocalize more when interacting with their infants in the lab than at home (Belsky, 1980). Further, Jaffe et al. (2001) report that the degree of novelty of the situation predicts the vocal activity (i.e., length of response latencies) on a gradient, meaning that in situations where infants interact with strangers in an unfamiliar environment they are expected to exhibit faster responses than when with strangers at home, which will exhibit faster responses than at home with their parent. These contradictory findings suggest that a closer look at turn-taking differences due to the familiarity of interlocutors and settings is needed.

## Activity

Analogously, different kinds of activities might involve different response latencies, as observed for other kinds of conversational behaviors (Dale, 2015; Dideriksen et al., 2020; Duran et al., 2019). While no within-study manipulation was reported in the literature, we can identify at least four different categories of activities during which response latencies were measured. (1) During free play, the adult and the child are generally recorded during spontaneously occurring play (Beebe et al., 1988; Hilbrink et al., 2015; Jaffe et al., 2001; Jasnow & Feldstein, 1986; Kelly, 1994; Kelly & Conture, 1992; Klann-Delius & Hofmeister, 1997; Northrup & Iverson, 2015; Savelkoul et al., 2007; Sawyer et al., 2017; Yaruss, 1997; Zlochower & Cohn, 1996). (2) During more structured everyday activities, the adult and child go through more routinized sequences of events, such as changing diapers (Leonardi et al., 2016; Malin et al., 2011; Marklund et al., 2015; Reissland & Stephenson, 1999). (3) During more focused conversations, the adult and child are focusing more specifically on the vocal interaction within structures at least partially imposed by the experimenter, for example, by providing a series of objects to be discussed (Bateson, 1975; Carvalho et al., 2019; Gratier et al., 2015; Kondaurova et al., 2020; Lange, 2016; Newman & Smit, 1989; Warlaumont et al., 2010, 2014; Welkowitz et al., 1990), and 4) highly constrained question-answer interactions (Casillas et al., 2016; Clark & Lindsey, 2015). No hypotheses concerning different latencies across these methods have been put forward in the literature.

## Language

Cross-linguistic variations observed by Stivers et al. (2009) in adult–adult conversations warrant a closer assessment of this factor in child-adult conversations. Indeed, one study of cross-linguistic variation in 2- to 5-month-old infants found significant differences in response-time latencies for three different languages (US English, French, Hindi; Gratier, 2003).

## Atypical groups

Finally, it might be insightful to study how turn-taking response time latencies vary within various atypical groups. Several studies point toward atypicalities in the development of turn-taking in a variety of different populations—both traditionally defined as atypical (e.g., autistic children) or as special (i.e., insecurely attached children). We refer to all of them as atypical in that they are argued to exhibit potentially atypical turn-taking. Children with cochlear implants as well as insecurely attached infants are slower at responding to their parents compared to more typically developing infants (Klann-Delius & Hofmeister, 1997; Kondaurova et al., 2020); prematurely born infants are faster (Reissland & Stephenson, 1999), and children with a stutter do not show any atypicality (Kelly, 1994; Kelly & Conture, 1992). On the other hand, parents of children with autism, and of prematurely born infants tend to have slower responses to their children, which might affect the child's responses (Reissland & Stephenson, 1999; Warlaumont et al., 2010, 2014). Furthermore, autistic individuals have often been argued to have conversational atypicalities and slower turn-taking dynamics (e.g., Choi et al., 2013; Heeman et al., 2010; Ochi et al., 2019; Ochs et al., 2004). However, no study so far has closely investigated which individual differences within the atypical groups might be related to turn-taking atypicalities, to better reveal which mechanisms are involved (see Weed et al., 2017 for a related example in autism spectrum disorder). From a theoretical perspective, identifying atypicalities in the development of turn-taking in a diverse range of populations might help us understand which mechanisms are indeed involved in turn-taking, for instance, if a group

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

presenting slower turn-taking also display a social, but not a linguistic impairment. Furthermore, given the likely foundational role of turn-taking in development, understanding turn-taking atypicalities might help us more effectively tackle linguistic and social irregularities in atypical groups (as in the social feedback loop hypothesis, Warlaumont et al., 2014).

## Summary

The reviewed literature highlights the importance of assessing how infants' cognitive and linguistic development, as well as the interpersonal dynamics inherent in conversations, affect the developmental trajectories of infants' turn-taking abilities. We should note, however, that the studies differ in the questions investigated, with the response latencies often only being reported as a side result. The studies also present conflicting results, perhaps due to confounds between factors (e.g., familiar interactors and free play as a method tend co-occur in most cases, which makes it difficult to disentangle their effects), or perhaps due to the heterogeneous settings, methods, and population samples investigated. In the following, we, therefore, provide a quantitative meta-analysis of the evidence available on the developmental trajectories of turn-taking and the possible factors there involved. We then complement it with a critical discussion of the limitations of the data and analysis, and provide a roadmap for future studies.

# MATERIALS AND METHODS

## Literature search

We adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Guidelines (PRISMA; Page et al., 2021) to ensure optimal transparency, as shown in Tables S1.1.1–1.1.2 and Figure S1.1. The full process was performed independently by two authors (VN and OV), as well as supervised and double checked by the two other authors (CC and RF). The systematic literature search was conducted on PubMed using the following search term: *"Early mother-infant communication" OR "protoconversation" OR "turn transition" OR "Turn alternation" OR "Turn offsets" OR "Turn timing" OR "Turn allocation" OR "Vocal responsiveness" OR "Vocal contingency" OR "Vocal congruence" OR "Turn-taking" OR "Conversational interaction" AND (infant OR child OR newborn OR baby OR babies)*. Neither the search, nor the subsequent meta-analysis were not pre-registered. As of March 2021, 232 possibly relevant studies were identified using this search term. In addition to using PubMed, we also performed a backwards literature search (assessing literature cited in our selected papers) and consulted Google Scholar to perform a forward literature search (assessing literature citing our selected papers), which resulted in the identification of an additional 34 studies.

The 266 studies obtained through this process were then screened using the following inclusion criteria: (i) studies should include dyadic vocal interactions between a child and an adult (therefore excluding, e.g., sign language, since different language modalities have been reported to unfold along different time scales, De Vos et al., 2015), (ii) studies should contain quantitative measures of temporal distance between turns (either adult to child or child to adult), (iii) children should be aged between 0 and 96 months. We opted for this 96-month cut-off because our main interest concerned the early development as well as the initial integration of the turn-taking structure and budding linguistic abilities. Both published and unpublished literature was considered for inclusion.

Two hundred and forty papers were accordingly excluded: 156 papers were excluded as unrelated to our topic (e.g., focused on physiological synchrony in the dyad, or alternating moves in a board game); 10 papers which did not include any vocal interaction (5 of which focused on sign language exchanges); 28 papers did not include any child–adult interactions; 46 papers did not contain quantitative measures of response latency. Several studies included multiple age groups or were of longitudinal nature, and therefore contained multiple response-time latencies. The final sample consisted of 26 unique papers and 78 response time latencies. Of these 26 papers, 12 included atypical groups, which constitute 22 of the 78 estimates of response-time latencies.

For more detailed information about the studies, the reader is referred to the Supporting Information. Specifically, Table S1.2 summarizes every study, including characteristics of the samples; Figure S1.3 shows a world map of the languages investigated (from exclusively North American and European sites); and Section S1.4 includes a discussion of possible biases in the search and selection of the studies.

## Citation network

To better explore the literature on infant-adult turn-taking, we extracted the reference list of each included paper using Web of Science, and filtered the lists to only include papers used in the meta-analysis. We used these data to build two different network plots, see Figure 2a,b. The first, a coupling network, identifies similarities in citing patterns and the discourses the papers refer to. In this network, two papers are connected if they both cite a common third paper. The more similar the bibliographical references are between the two papers, the stronger their connection. The second network, a direct-citation network, identifies which papers cite each other and the coherence among the contributions to the field. In both networks, we assessed the presence of clusters using the

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

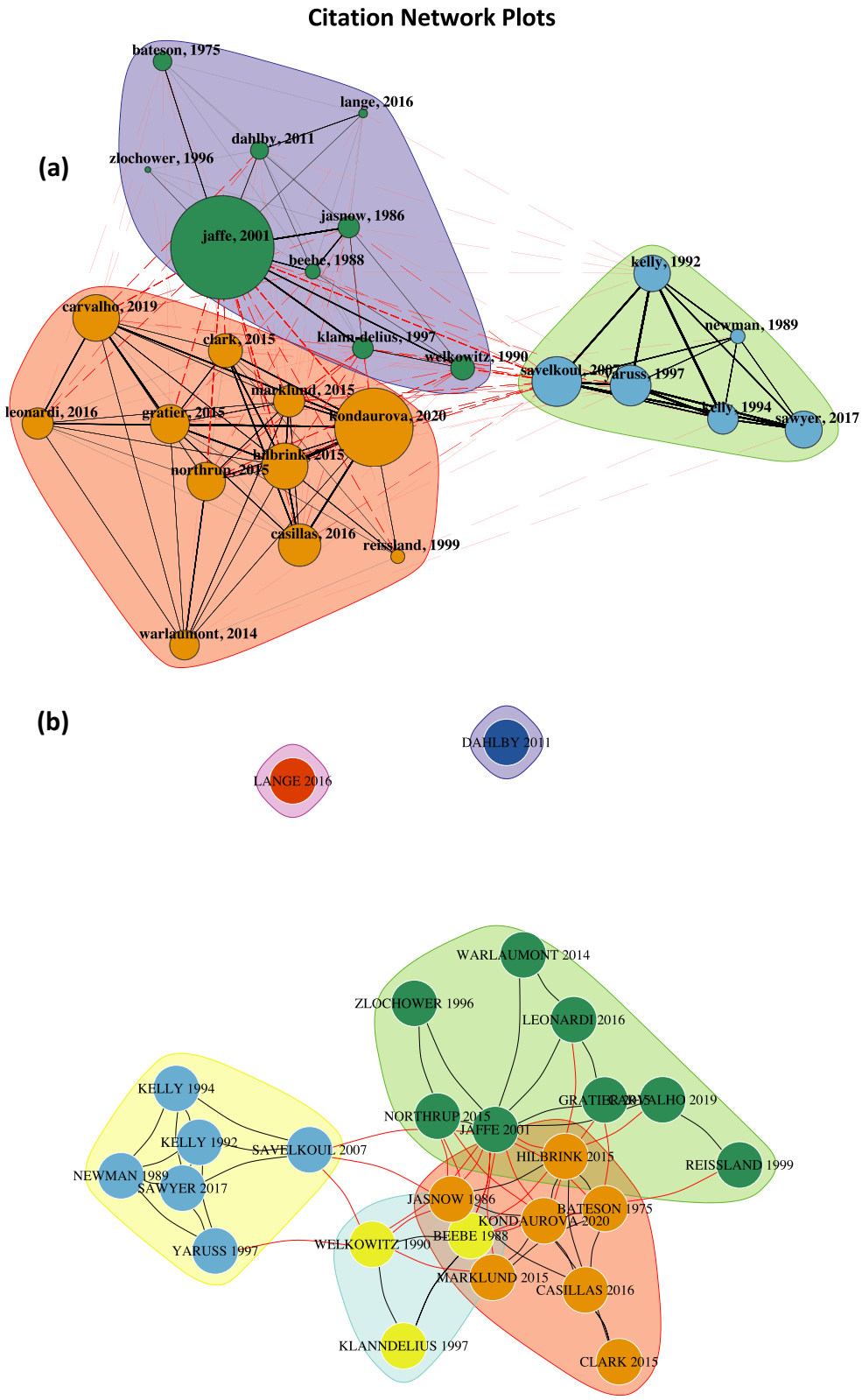


**FIGURE 2** Coupling and direct-citation networks for the studies reviewed. Plot a. above represents a coupling network, where papers are connected according to how much overlap there is in their bibliography. The size of the dots indicates the number and strength of links connected to that dot (paper). The below plot b represents a direct-citation network, where the papers are connected according to whether one of them cites the other. Colored clusters in both networks are inferred using a Louvain community detection algorithm and indicate papers more likely to be connected to each other than to papers outside of the cluster. The two isolates, Lange (2016) and Dahlby et al. (2011), represent grey literature, as they refer to a thesis and conference proceedings, respectively. In both panels, distance between two dots is an approximate measure of how similar (e.g., in connections to other dots) they are

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Louvain community algorithm, which identifies groups of papers (clusters) that are more likely to be connected to each other than to papers outside of the cluster (Traag & Bruggeman, 2009). To provide the reader with a point of comparison for citation network plots, we have added figure S2.1 from Cox et al. (2021) in Supporting Information. The visualizations relied on the *bibliometrix* and the *igraph* packages for R (Aria & Cuccurullo, 2017; Csardi & Nepusz, 2006).

In the coupling network in Figure 2a, the study populations and experimental conditions of the studies appear to be reflected in their citation patterns. The dense, blue cluster of studies, for example, primarily focuses on vocal interactions among atypically developing children and their parents (Kelly, 1994; Kelly & Conture, 1992; Newman & Smit, 1989; Savelkoul et al., 2007; Sawyer et al., 2017; Yaruss, 1997). The green cluster, on the other hand, is comprised of studies of typically-developing infants engaged in spontaneous interaction with their caregivers (Bateson, 1975; Beebe et al., 1988; Dahlby et al., 2011; Gratier et al., 2015; Jaffe et al., 2001; Jasnow & Feldstein, 1986; Klann-Delius & Hofmeister, 1997; Lange, 2016; Welkowitz et al., 1990; Zlochower & Cohn, 1996). Lastly, the orange cluster investigates child–caregiver interaction using more structured experimental activities, such as question-and-answer pairs (Carvalho et al., 2019; Casillas et al., 2016; Clark & Lindsey, 2015; Hilbrink et al., 2015; Kondaurova et al., 2020; Leonardi et al., 2016; Marklund et al., 2015; Northrup & Iverson, 2015; Reissland & Stephenson, 1999; Warlaumont et al., 2014). The direct-citation plot (Figure 2b) reflects a similar picture of a fragmented research field. For instance, the blue cluster of studies on atypical turn-taking exhibits a similar high degree of within-cluster coherence to that in Figure 2a. More generally, these citation plots indicate that the papers under investigation examine a wide variety of relevant aspects of infants' ability to take conversational turns, and that this manifests in the citation patterns of the studies.

## Data extraction

In order to perform a meta-analysis of the studies, we extracted the estimates for child and adult response latencies (mean and standard deviation) as well as available values of the following seven variables: (i) child age (mean and standard deviation), (ii) adult response latencies, (iii) familiarity of interlocutors and (iv) interactional settings, (v) type of activity, (vi) language being spoken, and (vii) whether the child had a developmental disorder or other relevant conditions. We further extracted recording duration and number of dyads involved in order to facilitate informed missing data imputation.

Response latencies were converted to milliseconds to ensure comparability and logarithmically transformed using the natural base. The transformation accounts for the long-tailed nature of the data, and was possible because no study reported negative average latencies (i.e., overlaps in vocalizations). Log-transformation is a non-linear transformation and that makes the interpretation of the model coefficients difficult. For instance, if an increase of 12 months in age were to relate to a decrease in response latency of 0.1 on a log scale, this increase is impossible to interpret unless we know the intercept. If the intercept (response latency when age is 0) is 7 on a log scale (about 1097 ms), a decrease of 0.1 implies a decrease of circa 104 ms (6.9 on a log scale being about 992 ms). If the intercept were 6 on a log scale (circa 403 ms), a decrease of 0.1 would only imply a decrease of 39 ms (5.9 on a log scale being about 365 ms). We, therefore, report effects of the moderators on the outcome scale to facilitate interpretation. In other words, we exponentiate the expected outcomes at different values of the moderators and report their differences. Age was converted to months, including two decimals, and scaled (subtracting the overall mean and dividing by the overall standard deviation) to facilitate model estimation.

Not all selected papers fully reported the relevant numeric estimates for our analyses. One paper (Kondaurova et al., 2020) only reported figures, from which we were able to extract the missing estimates using WebPlotDigitizer by Automeris. One paper (Hilbrink et al., 2015) only reported the median range of latencies, from which we extracted the average with the following formula: $\text{Average response latency} = \text{Max.} - \frac{\text{Max.} - \text{Min.}}{2}$. Given that the range covers the highest and lowest possible values, identifying the central value (half the distance between maximum and minimum values, counting down from the maximum value) provides a rough estimate of the mean with the assumption that the distribution is roughly symmetric.

Three papers only reported the child or the adult response latencies and eight papers only reported average estimates of response latencies, but not their variance; and these values could not be reconstructed from the papers. We did not contact the authors and chose to impute the missing data, using a multiple multivariate imputation process by chained equations as implemented in the R package *mice*, (Buuren & Groothuis-Oudshoorn, 2011). In order to account for the statistical uncertainty involved in the process of imputation (Azur et al., 2011; Sterne et al., 2009), the missing data were imputed 20 times, thus generating 20 separate datasets, using the following variables (where available): (i) child response latency (mean and standard deviation), (ii) adult response latency (mean and standard deviation), (iii) mean age of infants. Finally, these imputed values were post-processed to ensure realistic values: the ranges of the imputed values were restricted to the interquartile range of the observed data. All analyses in this paper were performed

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

on the resulting 20 datasets in which the missing data were imputed. We also tested the impact of the imputation, that is, we fitted models on the original and imputed datasets and ensured the results were analogous (see Supporting Information, S3). The raw data are available on OSF (https://osf.io/wkceb/?view_only=9cca387b49ef427fa9740cb94c3fbd5c) and will be made available on MetaLab (https://langcog.github.io/metalab/) upon publication of the manuscript.

Note that many additional dimensions are likely to impact turn-taking dynamics, but remain largely unreported in the studies: from ethnical and socioeconomic variation, to the more fine-grained structure of the interaction (e.g., was the sequence adult- or child-initiated?). These dimensions could not, therefore be systematically extracted and considered in this study.

## Statistical modeling

Meta-analytic models are usually deployed to pool the effect size estimates across the whole literature, weighted according to their variance, so as to achieve a more reliable estimate of the effect in the population. In this specific meta-analysis, we aimed at assessing the average child response latency in child-adult conversation and its relevant moderators. In other words, differently from most meta-analyses, our study is not concerned with aggregating effect sizes (e.g., differences between conditions, or existing correlation estimates and their uncertainty) from previous studies, since our questions of interest are mostly still unexplored. This study is concerned with aggregating the actual response latencies as they are reported, sometimes as secondary data. We then use these measures across studies to assess whether they are affected by possible moderators.

Since heterogeneous results are expected due to the differing nature of the methods and population samples in the studies, we relied on a random-effects model (Fernández-Castilla et al., 2020). By means of a multilevel structure, these models are able to address the heterogeneity between studies by assuming that every study and every sample has a different "true" response time latency, which does not necessarily overlap with the average response latency of the overall population. We modeled the outcome using a Student's t likelihood (i.e., modeling the error term of the model as following a Student's t distribution), where the standard deviation was given by the standard error of the effect size estimates, and the long-tails of the distribution allowed outliers to span out with only minimal effects on the estimates (Jylänki et al., 2011). This procedure produces robust inferential results, as single studies are weighted by the inverse of their variance (thus more uncertain effect estimates weigh less) and outliers exert a limited effect on the pooled response latency estimation (as they are accounted for in the long-tail of the *t*-distribution). Furthermore, a subset of studies reported trimming their data by excluding response latencies above a certain threshold (most commonly between 1 and 3 s). We explicitly modeled this censoring to make the estimates more comparable and fully account for available information. Estimates from the model are reported as mean and 95% compatibility intervals (CI) of the posterior estimates, that is, the range covered by the middle 95% of the posterior samples.

In order to investigate the developmental trajectory of turn-taking and whether the moderators we previously identified can systematically explain the variance in the response time latencies of infants across studies, we fitted three different models only on the typical population. The first model (*baseline model*) conditions the average response latency (on a log-scale) only to the variance due to articles and studies within articles.

Due to our specific interest in the developmental trajectory of response latency, our second model (*age model*) included the mean age of infants as a predictor. To account for the expected slow-down followed by a speed-up in turn-taking, we included age as a quadratic in order to appropriately model its effects. Age was also included in the random structure to explicitly model potential different effects of age in cross-sectional and longitudinal studies.

The third model (*moderators model*) included the following predictors: (i) mean age of the infants, (ii) logged adult response latency, (iii) familiarity of interlocutors, (iv) familiarity of interactional settings, (v) type of activity, and (vi) language. However, some contrasts might be confounded due to the sparsity of available data (e.g., data for non-English speaking infants coming from only one study with only familiar interlocutors). Therefore, we first checked cross-tabs in order to determine which contrasts should be tested and would allow us to make meaningful inferences (see Figure 3; Table 1). The resulting meaningful contrasts are reported in Table 2.

The subsequent models included also the data from atypical populations (i.e., autism, preterm, stutterers, insecurely attached, and hard of hearing), since systematic variations might provide cues as to which mechanisms modulate the development of turn-taking. The fourth and fifth models (*typical/atypical baseline and age models*) corresponded to the first and second models but we added a categorical variable which distinguished between the typical and five other atypical groups and added separate intercepts for these atypical groups.

The last model (*typical/atypical moderators model*) corresponds to the third moderators model but with a different intercept for every population and moderator. We planned to assess the impact of the same predictors as in model 3. However, when assessing the data available (see Table S9.1), we concluded that no meaningful between-group comparisons could be made, as it would be confounded by some conditions only happening in specific atypical groups. Therefore,

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

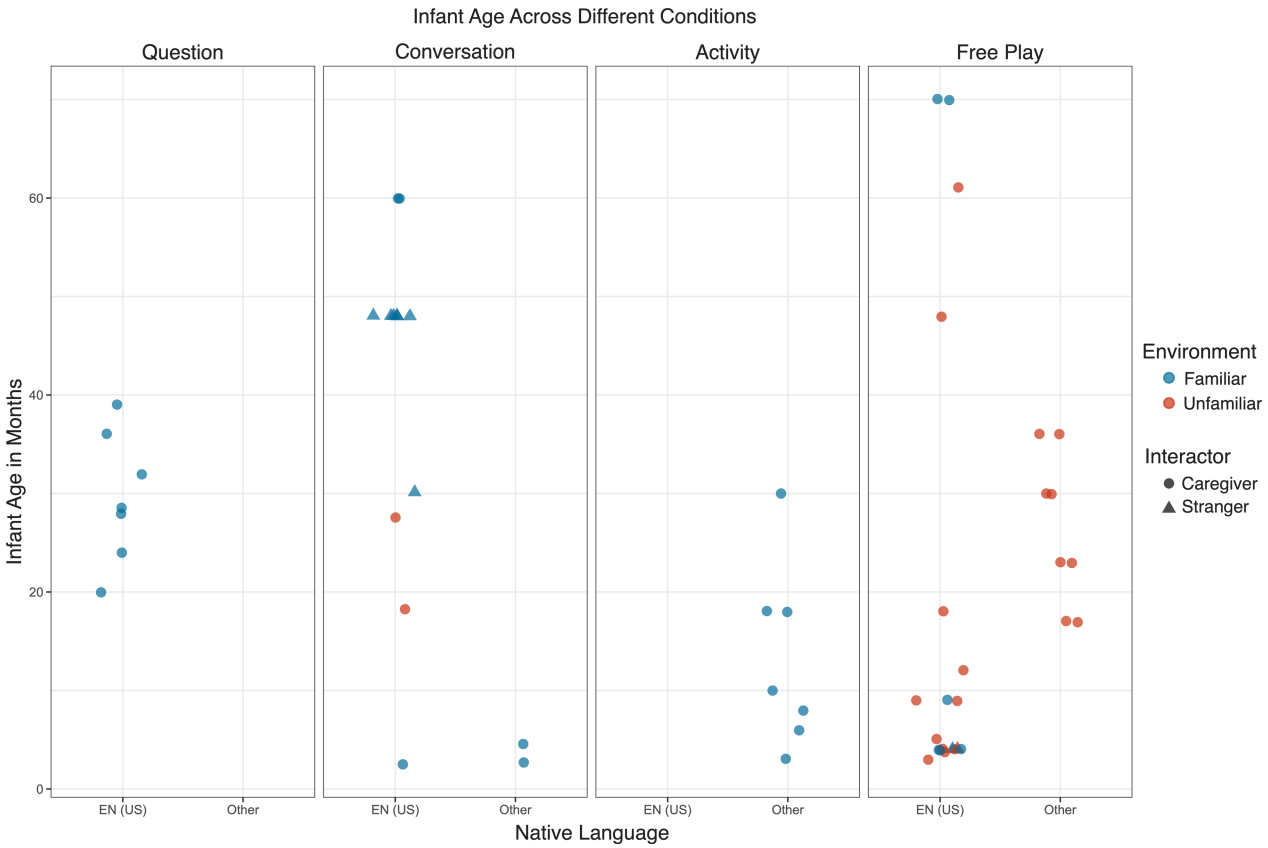


**FIGURE 3** Cross-tab showing the distribution of infant age across different activities, languages, environments and interactors

**TABLE 1** Cross-table for typical infants across method, environment, interactor, and language. Note. The abbreviation A under method represents "Activity," C represents "Conversation," P represents "Free play" and Q represents "question"

| | | | Method | | | |
|---|---|---|---|---|---|---|
| **Environment** | **Interactor** | **Language** | **A** | **C** | **P** | **Q** |
| Familiar | Caregiver | EN (US) | 0 | 3 | 6 | 7 |
| | | Other | 7 | 2 | 0 | 0 |
| | Stranger | EN (US) | 0 | 7 | 1 | 0 |
| | | Other | 0 | 0 | 0 | 0 |
| Unfamiliar | Caregiver | EN (US) | 0 | 2 | 11 | 0 |
| | | Other | 0 | 0 | 8 | 0 |
| | Stranger | EN (US) | 0 | 0 | 1 | 0 |
| | | Other | 0 | 0 | 0 | 0 |

it was decided that beyond the effect of age and adult response latency, the other planned moderators would not be examined.

Detailed description of the modeling choices, implementation, and quality checks are available in Supporting Information (see S3–S10). Moreover, we present exploratory analyses of how infant sex, country, and language affect our effects of interest in Figure S11.1 and Table S11.1.

**TABLE 2** Evidence ratios for moderators. See Table S5.1 for the detailed coefficients of each predictor

| **Moderator** | **Hypothesis tested** | **ER** |
|---|---|---|
| Age | Age <0 | 1.62 |
| $Age^2$ | $Age^2$ <0 | 8.77 |
| Adult response latency | Response latency >0 | >1000 |
| Environment | Familiar >unfamiliar | 1.45 |
| Method | Free play >conversation | 0.93 |
| Method | Question >conversation | 12.39 |
| Method | Question >free play | 22.93 |
| Language | Other >English | 3.35 |
| Interactor | Caregiver >stranger | 1.32 |

# RESULTS

## Response latency estimates

The baseline model was fitted to 55 measurements from 21 studies allowing us to estimate the overall intercept for infants' response latency. The overall estimate of infants' response latency was approximately 1 s (1089 ms [870, 1363]). If we include between-study variance, we can expect the mean response latency in a new study to range between 694 and 3559 ms (see Figure 4).

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

## Moderator Analysis for Heterogeneity

Including age (*age model*) provides an increased explained variance (from 14% to 25%, Bayesian $R^2$), but also a decrease in model generalizability (stacking weight = 0.053). Response latencies seem to increase with age up to 20 months of age, to then decrease (see Figure 5): At 1 month, the model predicted an infant response latency of 1289 ms [286, 4159], at 11 months of 1386 ms [547, 3148], at 20 months of 1429 ms [576, 3130], at 25 months of 1380 ms [522, 3128], at 30 months of 1307 ms [410, 3293], at 35 of 1251 ms [261, 3797]. However, caution is needed given the model is likely overfit (more explained variance, but lower generalizability). A control analysis including only conditions with a wide range of ages reported gave analogous results and is reported in the appendix (Figure S5.3).

Including additional moderators provided at least some improvement in generalizability compared to the age model (stacking weight = 0.614) as well as the baseline model (stacking weight = 0.581). In other words, its estimated generalizability to new data was comparable to that of the age model, but different enough predictions are produced that generalizability would be improved by averaging across the two models' predictions. The evidence available for how response latencies vary with most moderators seems fairly sparse and unreliable, with the exception of adults' response latency, and experimental activity of Question (cf. Table 2). While the infant response latency between studies exhibits heterogeneity,

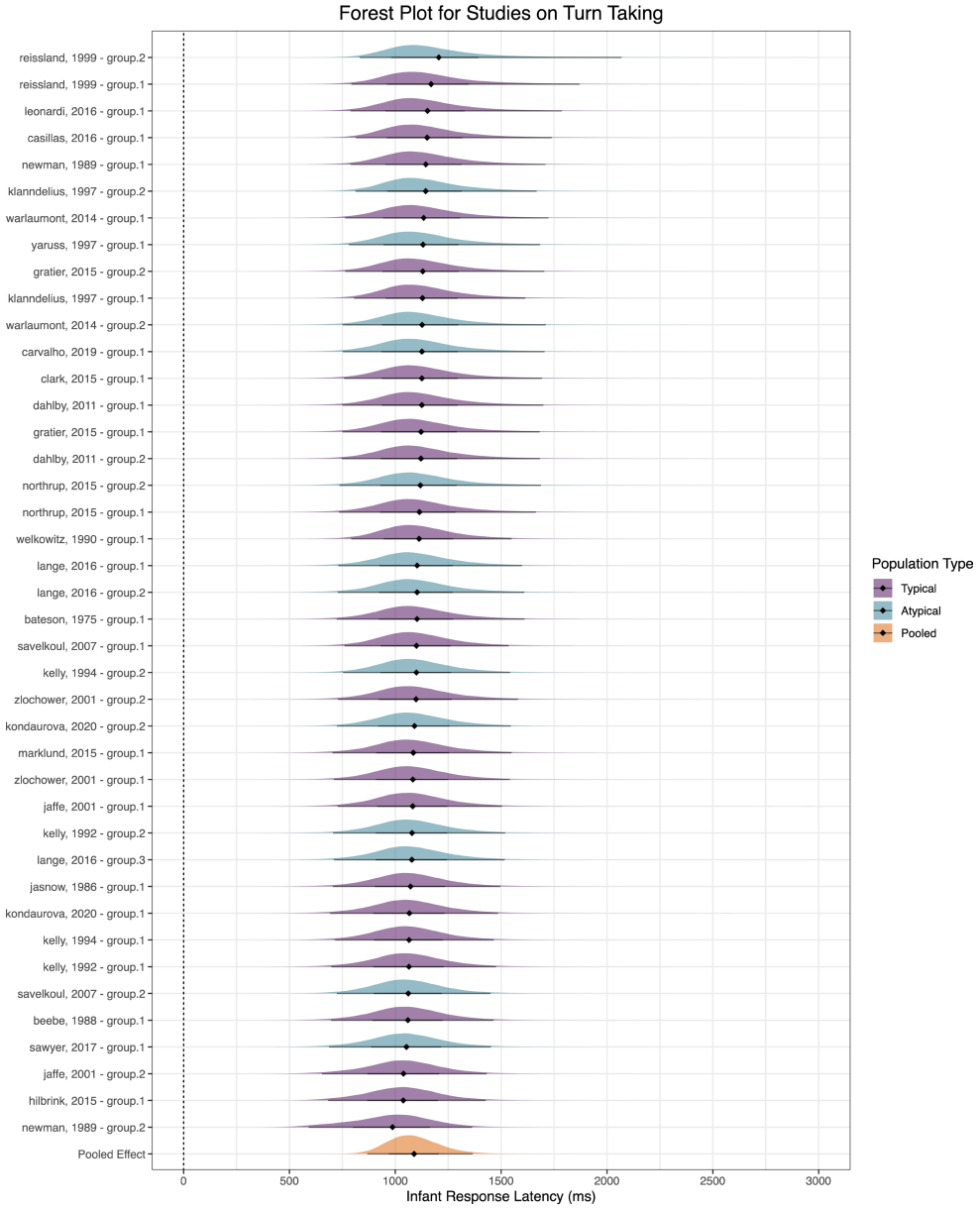


**FIGURE 4** Forest plot of the studies included in the meta-analysis. Studies of typical populations are indicated in black, and atypical in red. Note that the pooled effect estimate in the plot includes studies on atypical groups and therefore differs from that reported in the main text

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

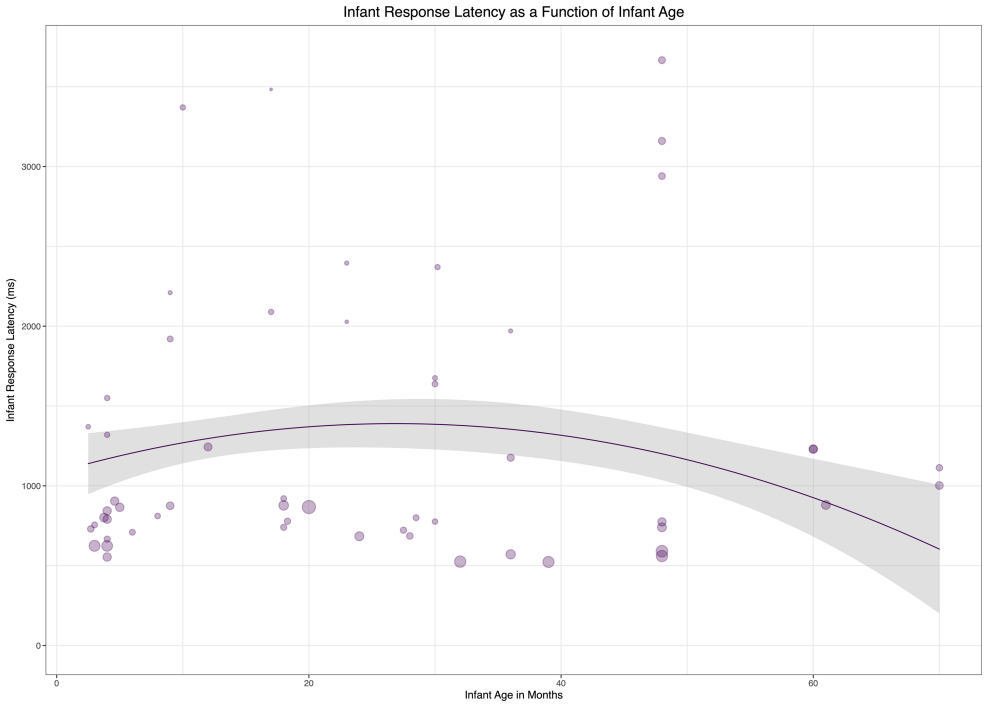


**FIGURE 5** Developmental trajectory of response time latencies: model prediction and actual data. The *x*-axis indicates the infants' average age in months, the *y*-axis response latencies in milliseconds. Red points represent the observed data, blue lines the model's inferred trajectory of response latencies, including shaded standard errors. The size of the points is inversely proportional to the standard error of the estimate (i.e., the larger the point, the smaller the standard error)

**TABLE 3** Overall estimates of the infants' response time latencies and range for the typical/atypical baseline model

| Population | Estimates [95% CIs] | Range including between study variability | Evidence ratio (compared to typical) |
|---|---|---|---|
| Autism | 1345 ms [528, 3003] | [212 ms, 11,403 ms] | Slower: 2.86 |
| Cochlear implant | 995 ms [445, 2060] | [284 ms, 6556 ms] | Faster: 1.00 |
| Typical | 1009 ms [806, 1295] | [604 ms, 3043 ms] | — |
| Insecurely attached | 1238 ms [676, 2200] | [504 ms, 6214 ms] | Slower: 2.97 |
| Preterm | 1411 ms [799, 2323] | [630 ms, 5521 ms] | Slower: 6.12 |
| Stutterer | 945 ms [637, 1404] | [462 ms, 3171 ms] | Faster: 1.62 |

they do seem to be robustly linked to those of adults ($\beta$ = 1.13 [0.78, 1.43] on a log scale). Keeping everything else constant, an increase of 1 s in adult response latency (e.g., 400 to 1400 ms) would be related to an increase of over 1 s in infant response latency (from 600 to 1857 ms). However, as adult response latency and infant age are correlated ($r$ = .33), caution in interpretation should be exercised. Within the meaningfully comparable contrasts (cf. Table 1; Figure 3), most of the moderators show no reliable influence on infant response latencies (evidence ratio [ER] <2 and >.5), with the exception of Question as experimental activity, which elicits slower infant response latencies than the other activities (ER >10).

## Typical/atypical baseline model

The *typical/atypical baseline model* was fitted to 78 measurements from 26 studies allowing us to estimate and compare the separate intercepts for the typical and different atypical groups (see Table 3). Autistic, preterm, and insecurely attached children exhibited slower response latencies than typically developing children. The typical/atypical baseline model was able to explain 21% of the observed variance.

## Typical/atypical moderator models

Including *age* as a predictor of infant response latency did not provide additional generalizability (stacking weight = 0.00), and decreased explained variance (from 21% to 9%). However, the expected effects of age within each population were rather small and uncertain, due to the sparsity of data (see Figure 6).

Adding *adult response latency* to the age model resulted in increased generalizability (stacking weight = 0.879) and explained variance (from 9% to 44%). We observe

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that adult response latency is generally the more robust predictor (except for insecurely attached infants), while age becomes more uncertain (see Table 4). However, as the data are sparse (see Figure S9.1), much interpretive caution should be exercised.

## DISCUSSION

The current study aimed at better understanding the current state of research on the development of turn-taking behaviors in human infants. In particular, we wanted to investigate the developmental trajectory of infants' turn-taking abilities and identify the key moderators affecting them in order to better constrain the space of potential mechanisms underlying turn-taking.

The above meta-analysis indicates that typically developing children exhibit an average response latency of about 1s and that atypically developing infants exhibit slower response latencies (i.e., particularly preterm infants and infants with autism and insecure attachment). The meta-analysis also provides information about the influence of moderators. Specifically, the results indicate that infant response latencies might exhibit a quadratic trajectory of development in early infancy and an increase until 20 months of age. Furthermore, these latencies exhibit a strong correlation with adult response latencies in both typical and atypical infants. This correlation highlights the interactional nature of turn-taking, as discussed further below. The analysis also indicates that insufficient evidence is available as to how infant response latencies vary according to age, language, methodology, and the familiarity of the environment and interactor. In the following, we contextualize these findings in the wider literature and comment on developmental trajectories, interpersonal dimensions, atypical development, and future directions of research on turn-taking.

### Developmental trajectory of turn-taking

A popular model (the Interaction Engine Hypothesis, introduced in Levinson, 2006) of turn-taking suggests a slowdown at around 12 months of age as children integrate their interactional abilities with their growing capacities for language (i.e., their intentional use of words in particular), followed by a speed-up at around 18 months. Earlier studies place the slowdown at 9 months (Hilbrink et al., 2015), or even 4–5 months of age (Gratier et al., 2015), and do not investigate the subsequent speed-up. The meta-analytic pattern is very uncertain and might point to an early slowdown (Gratier et al., 2015) and a speed-up at around 20 months of age. Albeit cautiously, due to the aggregated data analyzed and likely overfitting of the model, we speculate that this pattern is not in contradiction with the Interaction Engine Hypothesis. However, the continuous nature of the processes underlying language acquisition should play a more important role. In particular, the growing

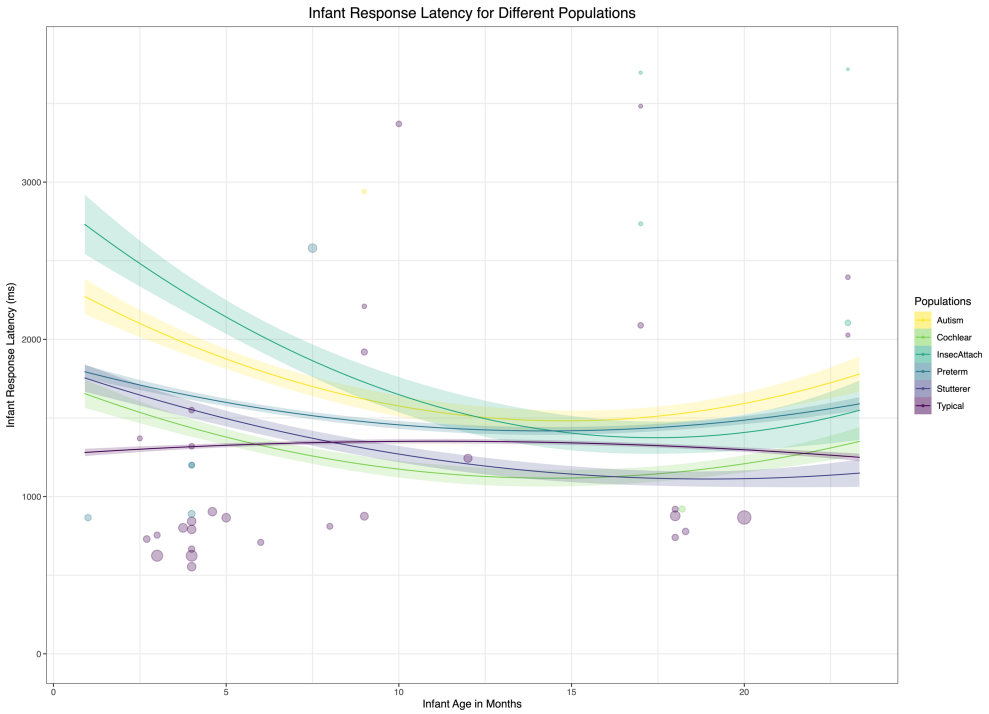


**FIGURE 6** Estimated effects of age in typical and atypical populations. The shaded area indicates the standard error of the estimated population mean for each given group. The size of the points is inversely proportional to the standard error of the estimate (i.e., the larger the point, the smaller the standard error)

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

recognition of the importance of intentional communication, and the "9-month revolution" (Tomasello, 1999, as suggested in Hilbrink et al., 2015) might not be the one determining factor for response time latencies of infants, or at least should have a much more diffuse and gradual impact. Furthermore, one should more closely consider the content, complexity, and contexts of interactions and turn-by-turn exchanges, as we might expect them to affect response latencies. For instance, Casillas et al. (2016) have shown that the complexity of questions differentially affect children at different ages, with complex questions being answered relatively more slowly by younger children. The meta-analytic results likewise provide evidence for questions exhibiting slower response time latencies (see Table 2). As meta-analytic data could present multiple confounds, however, these findings emphasize the importance of longitudinal studies including data on both the linguistic and social development of the child, so we can better assess the concurrent development of these measures and response time latencies. This could also help to build more detailed versions of the Interaction Engine Hypothesis and to generate more specific predictions about how and why trajectories might differ on a child-to-child basis.

In any case, the response latencies we found (roughly 1 s) are much slower than those reported for adult–adult conversations (ranging between 7 and 470 ms, see Stivers et al., 2009). This result suggests that even the older children in our data (8-year-olds) fail to display clear signs of anticipating conversational turns in their response latencies and that their developmental trajectories are far

**TABLE 4** Evidence ratios for the moderator effects in the different populations

| Population | Hypothesis tested | ER |
|---|---|---|
| Autism | Age <0 | 1.02 |
| | $Age^2$ >0 | 1.09 |
| | Adult response latency >0 | 1.68 |
| Cochlear implant | Age >0 | 1.08 |
| | $Age^2$ <0 | 1.04 |
| | Adult response latency >0 | 1.14 |
| Typical | Age <0 | 2.31 |
| | $Age^2$ <0 | 11.81 |
| | Adult response latency >0 | >1000 |
| Insecure attached | Age <0 | 1.56 |
| | $Age^2$ >0 | 1.09 |
| | Adult response latency <0 | 1.08 |
| Preterm | Age <0 | 1.23 |
| | $Age^2$ >0 | 1.15 |
| | Adult response latency >0 | 3.92 |
| Stutterer | Age <0 | 1.16 |
| | $Age^2$ <0 | 1.56 |
| | Adult response latency >0 | 1.25 |

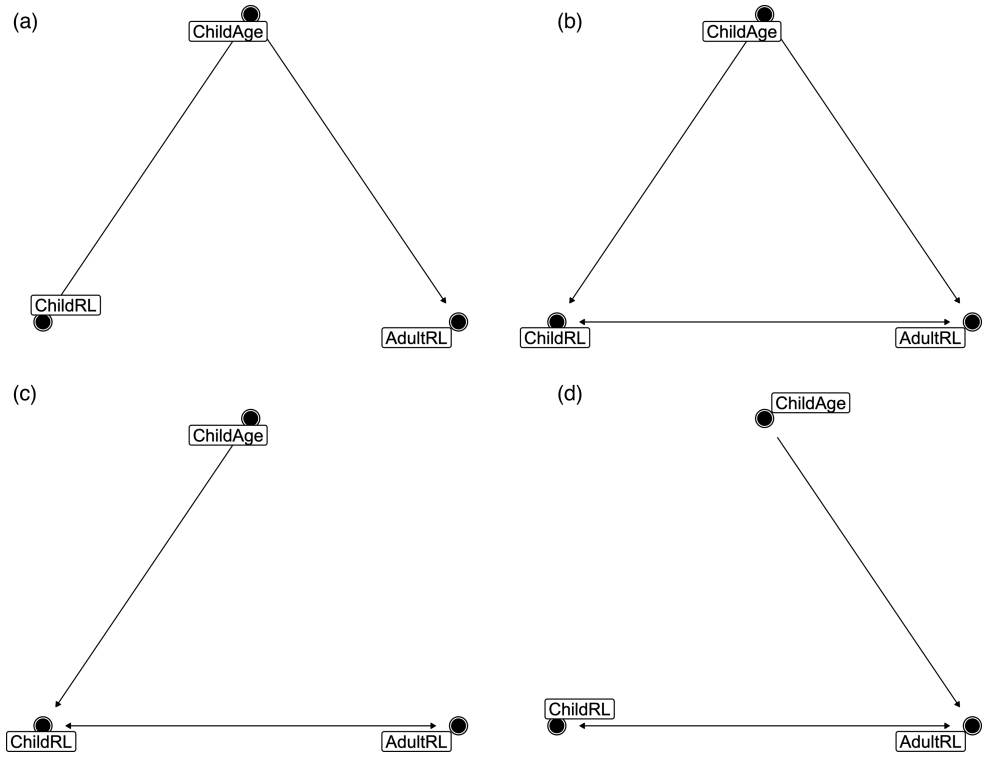


**FIGURE 7** Potential causal structures underlying the dependency among child age, child response latencies, and adult response latencies. Each node represents a variable, and each arrow a hypothesized causal relation. Note that in contrast to conventions in the visualization of causal connections, here we occasionally allow for bidirectional causality: infants and adults influencing each other

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from complete. On the other hand, previous work has shown that 3- to 5-year-old children anticipate conversational turns in a controlled task, despite the longer latencies, which might be due to language production skills still being in development (Lindsay et al., 2019).

## Interpersonal adaptation

Turn-taking has often been conceptualized as a shared, cooperative process (Fusaroli, Rączaszek-Leonardi, et al., 2014; Manson et al., 2013; Pickering & Garrod, 2021; Wilson & Wilson, 2005). Interlocutors come to develop a shared tempo in their syllable lengths and response latencies, thus making predictions and alternation in speakers smoother. In line with these ideas, we observe a strong relation between the response latencies of children and adult interlocutors (Figure 8). Given the coarse graining of the data we have available (summary statistics at the study level), it is hard to assess whether both children and adults adapt to each other and whether this happens at a turn-by-turn basis or at a slower timescale. Therefore, one can but speculate as to whether adults are simply adjusting to the longer time required by children to answer as they engage in more and more complex conversations, or are more actively providing scaffolding to the children's development, implicitly helping them to take enough time (Gratier et al., 2015; Hilbrink et al., 2015; Stern et al., 1977). In Figure 7 below, we depict four potential causal structures underlying the interrelatedness among child age, child response latency, and adult response latency. Structures B, C, and D would apply if the developmental trajectory of infant response latencies depends in part on their relation to caregivers' response time latencies. This process of mutual adaptation may in turn be affected by child age in various plausible ways. For example, (i) child age may primarily affect children's response time latencies and adults may then adapt the timing of their own responses accordingly (structure C). Alternatively, (ii) if adults adjust their response time latencies to scaffold their child's development, child age may primarily affect adult response time latencies, to which children then adapt their own response latencies (structure D). Finally, (iii) child age could also influence both child and adult response latencies, and child-adult adaptation may take place to varying extents (structures A & B).

In order to better understand the interpersonal dynamics of turn-taking and the reciprocal roles of child and adult response latency, we need both experimental setups manipulating response latencies (Newman & Smit, 1989) as well as fine-grained longitudinal datasets, with turn-by-turn latencies (Weed et al., 2017). Availability of turn-by-turn data would allow more nuanced analysis of how self-regulation (maintaining a certain average tempo) and interpersonal adaptation (following the other's tempo) of responses unfold in real time (Dale et al., 2013; Ferrer & Helm, 2013; Fusaroli et al., 2016; Rast & Ferrer, 2018; Takahashi et al., 2013).

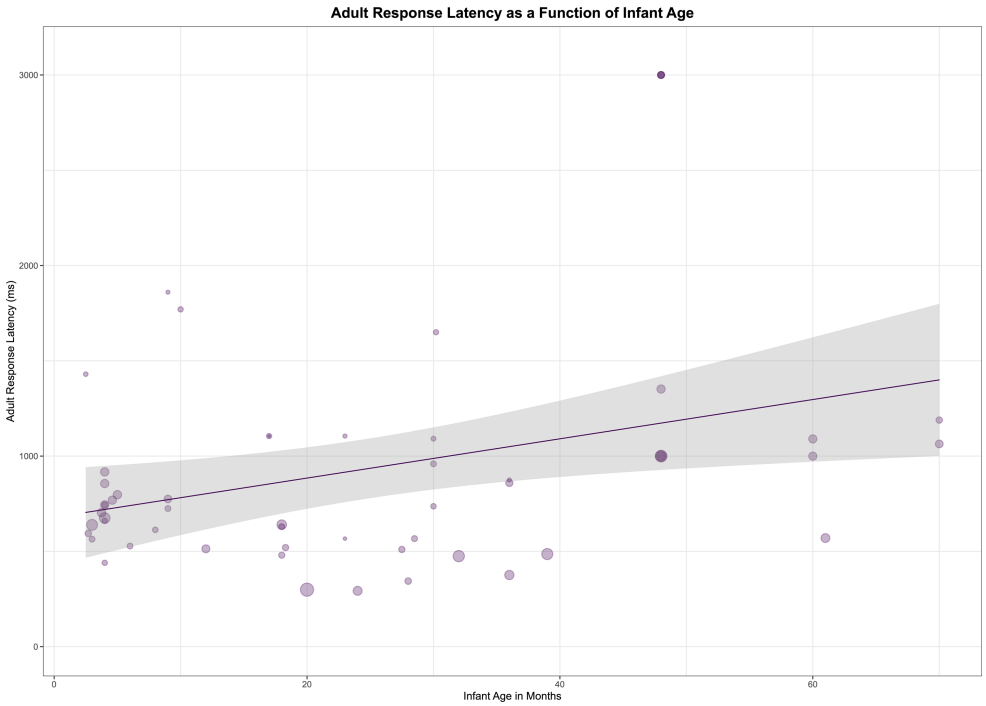


**FIGURE 8** Plot of adult response latency as a function of infant age for the data on typical infants. The plot shows a positive correlation between the two moderator variables

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

## Typical/atypical populations and other moderators

We included atypical populations with a diverse range of impairments in this meta-analysis, with a view to investigating how these conditions affect response latencies and the developmental trajectories of infants' turn-taking abilities. Preterm infants and infants diagnosed with autism and insecure attachment stood out as having the slowest response latencies. While these groups involve both social and linguistic developmental impairment, which might motivate the slower turn-taking patterns (Klann-Delius & Hofmeister, 1997; Warlaumont et al., 2014), the sparsity of the data make mechanistic explanations untenable. Although the data were too sparse to provide realistic descriptions of the developmental trajectories in these groups, the connection between infant and adult response latencies seems to be held constant across most of the groups.

We do not observe any reliable effects from the familiarity of environment and interactors, method, and language, with the exception of the experimental activity of Question exhibiting slower infant response latencies. These null effects and the finding of an influence of constrained question-answer interactions (Casillas et al., 2016; Clark & Lindsey, 2015) on infant response latencies should be interpreted with caution. The data available to contrast these conditions remain sparse, are often confounded across moderators, and aggregate studies with substantial differences within the same condition (e.g., studies focusing only on questions-answers and other studies including all response latencies). Targeted studies with within-subject manipulations are therefore needed to investigate these more subtle effects, for instance by recording the same infants interacting with familiar and unfamiliar interlocutors in familiar and unfamiliar environments.

## Future directions

Systematic reviews and meta-analyses function as moments of self-reflection for a field of research, more aimed at identifying good practices, pitfalls, and new directions than state-of-the-art estimates, which are likely inaccurate (see, Fusaroli et al., 2022; Lewis et al., 2020).

Our exploratory network analyses indicated a field examining a wide variety of aspects of the development of turn-taking, often reporting infant response latencies as side results to more central research questions. We intend for this systematic review and meta-analysis to serve as common ground among researchers interested in the development of turn-taking as well as in neighboring fields of animal communication, adult–adult interactions, and human–computer interactions.

Furthermore, the field could benefit from the development of shared curated resources, both data and computational models. Most of the studies included in the meta-analysis presented several issues for making strong inferences: (i) they were cross-sectional, or covered only short longitudinal timespans, (ii) they presented roughly aggregated data (e.g., age instead of measures of linguistic and social development); (iii) they provided group-level summary statistics instead of individual-level data; (iv) they contributed interaction-level estimates, instead of turn-by turn response latencies; and (v) the data lacked variability: 51 of the 78 effect sizes reported concerned US English (see also Figure S1.3 for a representation of the geographical reach of the data). In order to advance our understanding of the development of turn-taking, we need access to comprehensive cross-linguistic and more representative (e.g., of socioeconomic, cultural, and ethnic diversity) datasets that provide turn-by-turn response latencies and the possibility to code for conversational moves (Bergey et al., 2021; Nikolaus et al., 2021), in combination with longitudinal assessments of linguistic and social development (Fusaroli et al., 2019; Naigles & Fein, 2017). A recent paper and R package has also been produced to provide streamlined standard ways of preprocessing the data (Casillas & Scaff, 2021), and other approaches are being developed to identify diverse types of interactional sequences and provide a partially automated coding of them (Bergey et al., 2021; Nikolaus et al., 2021).

Better shared data resources can provide the necessary testing ground for computational models of turn-taking and turn-taking development. We have mentioned the Interaction Engine Hypothesis, Coupled Oscillators models, and dyadic models of self-regulation and adaptation, which are promising but currently underspecified to capture the key patterns we see in the meta-analytic data. By providing computational implementations of these and other models, future research can be more explicit about the assumptions of the theories, as well as compare the predictions of different models and test them on the data available (Guest & Martin, 2021; Navarro, 2021; Rich et al., 2021). Furthermore, the development of better mechanistic models can help the field better capture the meaningful aspects of cross-linguistic, sociodemographic, ethnic, and cultural diversity. Better articulating which aspects of diversity might play a role in the mechanisms at stake, we can create more representative samples for those aspects and develop more meaningful causal inference analyses (Christiansen et al., 2022; Deffner et al., 2021).

Finally, our understanding of the development of turn-taking could benefit from a more explicit dialogue with the study of turn-taking across language modalities (e.g., sign languages and co-speech gestures), adult–adult turn-taking as well as animal communication. Different language modalities have been reported to unfold at different time scales (De Vos et al., 2015), and could bring important discussions about how timing is adjusted to different contextual constraints. Adult–adult studies can

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

provide a picture of fully developed turn-taking abilities and a target for child–adult trajectories, while animal communication studies can provide both a comparative and evolutionary perspective (Demartsev et al., 2018; Pika et al., 2018; Takahashi et al., 2013, 2016).

## CONCLUSION

This systematic review and meta-analysis revealed a possible developmental trajectory of infant response latencies: there is uncertain evidence that response latencies might gradually increase up to 20 months, and quite reliable evidence that they exhibit a strong relation to the adult interlocutors' response latency. Based on these results, we provide recommendations and suggest directions for future studies. We advocate for the development of shared longitudinal diverse corpora with turn-by-turn data and comprehensive assessment of infants' linguistic and social development. We also suggest more explicit definition and testing of computational models in order to inform the cognitive mechanisms underlying turn-taking.

### ETHICS STATEMENT

The manuscript relies on publicly available data (published articles) and has been deemed exempt from the need of ethical approval by the local ethical committee.

### ACKNOWLEDGMENTS

VN and OV were funded by an Erasmus Internship program. RF was funded by the NIH-NIDCD R01 project "Early Predictors to School Age Language" and seed funding from the Interacting Minds Center.

### CONFLICT OF INTEREST

RF has been a paid consultant for F. Hoffman La Roche on related but non-overlapping topics.

### DATA AVAILABILITY STATEMENT

The raw data are available on OSF (https://osf.io/wkceb/?view_only=9b90de634b9d4defad8da4e04ede7068) and will be made available on MetaLab (https://langcog.github.io/metalab/) upon publication of the manuscript.

### ORCID

*Otto Versyp* https://orcid.org/0000-0002-9536-3634
*Riccardo Fusaroli* https://orcid.org/0000-0003-4775-5219

### REFERENCES

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*(1), 40–49. https://doi.org/10.1002/mpr.329

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

Belsky, J. (1980). Mother-infant interaction at home and in the laboratory: A comparative study. *The Journal of Genetic Psychology*, *137*(1), 37–47. https://doi.org/10.1080/00221325.1980.10532800

Bergey, C., Marshall, Z., DeDeo, S., & Yurovsky, D. (2021). Learning communicative acts in children's conversations: A hidden topic Markov model analysis of the CHILDES Corpora. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12591

Bigelow, A. E. (1998). Infants' sensitivity to familiar imperfect contingencies in social interaction. *Infant Behavior and Development*, *21*(1), 149–162. https://doi.org/10.1016/S0163-6383(98)90060-1

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

Cappella, J. N. (1996). Dynamic coordination of vocal and kinesic behavior in dyadic interaction: Methods, problems, and interpersonal outcomes. In J. H. Watt & C. A. VanLear (Eds.), *Dynamic patterns in communication processes* (pp. 353–386). Sage Publications Inc.

Carvalho, M. E. S., Justo, J. M. R. M., Gratier, M., Tome, T., Pereira, E., & Rodrigues, H. (2019). Vocal responsiveness of preterm infants to maternal infant-directed speaking and singing during skin-to-skin contact (Kangaroo Care) in the NICU. *Infant Behavior and Development*, *57*. https://doi.org/10.1016/j.infbeh.2019.101332

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 annual meeting of the Cognitive Science Society*, 35(35). https://escholarship.org/uc/item/9qz9m12w

Casillas, M., & Scaff, C. (2021). Analyzing contingent interactions in R with 'chattr'. *Proceedings of the annual meeting of the Cognitive Science Society*, 43(43). https://escholarship.org/uc/item/4rr848x0

Choi, J., Lee, Y., Choi, J., & Lee, Y. (2013). Conversational turn-taking and topic manipulation skills of children with high-functioning autism spectrum disorders. *Communication Sciences & Disorders*, *18*(1), 12–23. https://doi.org/10.12963/csd.13002

Christiansen, M., Contreras, P., & Trecca, F. (2022). Towards a comparative approach to language acquisition. *Current Directions in Psychological Science*. https://doi.org/10.1177/09637214211049229

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*, https://doi.org/10.3389/fpsyg.2015.00890

Cox, C. M. M., Keren-Portnoy, T., Roepstorff, A., & Fusaroli, R. (2021). A Bayesian meta-analysis of infants' ability to perceive audio–visual congruence for speech. *Infancy*. https://doi.org/10.1111/infa.12436

Cross, L., Michael, J., Wilsdon, L., Henson, A., & Atherton, G. (2020). Still want to help? Interpersonal coordination's effects on helping behaviour after a 24 hour delay. *Acta Psychologica*, *206*, 103062. https://doi.org/10.1016/j.actpsy.2020.103062

Crown, C. L. (1991). Coordinated interpersonal timing of vision and voice as a function of interpersonal attraction. *Journal of Language and Social Psychology*, *10*(1), 29–46. https://doi.org/10.1177/0261927X91101002

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Cummins, F. (2013). Towards an enactive account of action: Speaking and joint speaking as exemplary domains. *Adaptive Behavior*, *21*(3), 178–186. https://doi.org/10.1177/1059712313483144

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Curioni, A., Vesper, C., Knoblich, G., & Sebanz, N. (2019). Reciprocal information flow and role distribution support joint action coordination. *Cognition*, *187*, 21–31. https://doi.org/10.1016/j.cognition.2019.02.006

Dahlby, M., Irmalm, L., Kytöharju, S., Wallander, L., Zachariassen, H., Ericsson, A., & Marklund, U. (2011). Parent-child interaction: Relationship between pause duration and infant vocabulary at 18 months. *Fonetik 2011*, *1*, 101–104.

Dale, R. (2015). An integrative research strategy for exploring synergies in natural language performance. *Ecological Psychology*, *27*(3), 190–201. https://doi.org/10.1080/10407413.2015.1068649

Dale, R., Fusaroli, R., Duran, N. D., & Richardson, D. C. (2013). The self-organization of human interaction. *Psychology of Learning and Motivation*, *59*, 43–95. https://doi.org/10.1016/B978-0-12-407187-2.00002-2

de Reus, K., Masayo, S., Marianna, A., Gamba, M., de Heer Kloots, M., Miriam, L., Bruno, J. H., Laurel, T., & Ravignani, A. (2021). Rhythm in dyadic interactions. *PsyArXiv*. https://doi.org/10.31234/osf.io/9yrkv

De Vos, C., Casillas, M., Crasborn, O., & Levinson, S. C. (2015). *Is turn timing dependent on language modality?* The 36th TABU Dag.

Deffner, D., Rohrer, J. M., & McElreath, R. (2021). A causal framework for cross-cultural generalizability. *PsyArXiv*. https://doi.org/10.31234/osf.io/fqukp

Demartsev, V., Strandburg-Peshkin, A., Ruffner, M., & Manser, M. (2018). Vocal turn-taking in meerkat group calling sessions. *Current Biology*, *28*(22), 3661–3666.e3. https://doi.org/10.1016/j.cub.2018.09.065

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2020). Quantifying the interplay of conversational devices in building mutual understanding. *PsyArXiv*. https://doi.org/10.31234/osf.io/a5r74

Dominguez, S., Devouche, E., Apter, G., & Gratier, M. (2016). The roots of turn-taking in the neonatal period. *Infant and Child Development*, *25*(3), 240–255.

Dressler, R. A., Buder, E. H., & Cannito, M. P. (2009). Rhythmic patterns during conversational repairs in speakers with aphasia. *Aphasiology*, *23*(6), 731–748. https://doi.org/10.1080/02687030802165582

Duran, N. D., & Fusaroli, R. (2017). Conversing with a devil's advocate: Interpersonal coordination in deception and disagreement. *PLoS One*, *12*(6), e0178140. https://doi.org/10.1371/journal.pone.0178140

Duran, N. D., Paxton, A., & Fusaroli, R. (2019). ALIGN: Analyzing linguistic interactions with generalizable techNiques—A Python library. *Psychological Methods*, *24*(4), 419–438. https://doi.org/10.1037/met0000206

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*, *52*(5). https://doi.org/10.3758/s13428-020-01373-9

Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology*, *88*(3), 296–308. https://doi.org/10.1016/j.ijpsycho.2012.10.013

Fusaroli, R., Bjørndahl, J. S., Roepstorff, A., & Tylén, K. (2016). A heart for interaction: Shared physiological dynamics and behavioral coordination in a collective, creative construction task. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(9), 1297.

Fusaroli, R., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., & Weed, E. (2022). Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Research*, https://doi.org/10.1002/aur.2661

Fusaroli, R., Konvalinka, I., & Wallot, S. (2014). Analyzing Social Interactions: The Promises and Challenges of Using Cross Recurrence Quantification Analysis. *Springer Proceedings in Mathematics & Statistics*, *103*, 137–155. https://doi.org/10.1007/978-3-319-09531-8_9

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157. https://doi.org/10.1016/j.newideapsych.2013.03.005

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

Fusaroli, R., Weed, E., Fein, D., & Naigles, L. (2019). Hearing me hearing you: Reciprocal effects between child and parent language in autism and typical development. *Cognition*, *183*, 1–18. https://doi.org/10.1016/j.cognition.2018.10.022

Fusaroli, R., Weed, E., Fein, D., & Naigles, L. (2021). Caregiver linguistic alignment to autistic and typically developing children. *PsyArXiv*. https://doi.org/10.31234/osf.io/ysjec

Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, *15*(2), 221–235. https://doi.org/10.1017/S0305000900012344

Gratier, M. (2003). Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience. *Cognitive Development*, *18*(4), 533–554. https://doi.org/10.1016/j.cogdev.2003.09.009

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpgyg.2015.01167

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279. https://doi.org/10.1111/1467-9280.00255

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, *16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Heeman, P. A., Lunsford, R., Selfridge, E., Black, L., & van Santen, J. (2010). Autism and interactional aspects of dialogue. *Proceedings of the 11th annual meeting of the Special Interest Group on Discourse and Dialogue* (pp. 249–252).

Heinemann, T. (2010). The question–response system of Danish. *Journal of Pragmatics*, *42*(10), 2703–2725. https://doi.org/10.1016/j.pragma.2010.04.007

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*. https://doi.org/10.3389/fpsyg.2015.01492

Hsu, H.-C., & Fogel, A. (2003). Stability and transitions in mother-infant face-to-face communication during the first 6 months: A microhistorical approach. *Developmental Psychology*, *39*(6), 1061–1082. https://doi.org/10.1037/0012-1649.39.6.1061

Hsu, H.-C., Fogel, A., & Messinger, D. S. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction: Factors associated with quantitative and qualitative differences. *Infant Behavior & Development*, *24*(1), 107–128. https://doi.org/10.1016/S0163-6383(01)00061-3

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 255. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Jaffe, J., Beebe, B., Feldstein, S., Crown, C., & Jasnow, M. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*, i–149.

Jasnow, M., & Feldstein, S. (1986). Adult-like temporal characteristics of mother-infant vocal interactions. *Child Development*, *57*(3), 754–761. https://doi.org/10.2307/1130352

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Jylänki, P., Vanhatalo, J., & Vehtari, A. (2011). Robust Gaussian process regression with a student-t likelihood. *Journal of Machine Learning Research*, *12*(11), 31.

Kaye, K. (1977). Toward the origin of dialogue. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 89–117). Academic Press.

Kelly, E. (1994). Speech rates and turn taking behaviors of children who stutter and their fathers. *Journal of Speech and Hearing Research*, *37*(6), 1284–1294. https://doi.org/10.1044/jshr.3706.1284

Kelly, E., & Conture, E. (1992). Speaking rates, response-time latencies and interrupting behaviors of young stutterers, nonstutterers, and their mothers. *Journal of Speech and Hearing Research*, *35*(6), 1256–1267. https://doi.org/10.1044/jshr.3506.1256

Klann-Delius, G., & Hofmeister, C. (1997). The development of communicative competence of securely and insecurely attached children in interactions with their mothers. *Journal of Psycholinguistic Research*, *26*(1), 69–88. https://doi.org/10.1023/a:1025012221407

Kondaurova, M. V., Smith, N. A., Zheng, Q., Reed, J., & Fagan, M. K. (2020). Vocal turn-taking between mothers and their children with cochlear implants. *Ear and Hearing*, *41*(2), 362–373. https://doi.org/10.1097/AUD.0000000000000769

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., Orden, G. V., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8514–8519. https://doi.org/10.1073/pnas.1016955108

Lange, A. E. (2016). *Vocal rhythm coordination and preterm infants: Rhythms of dialogue in a high-risk NICU sample* (Vol. *1*). City University of New York.

Leonardi, G., Nomikou, I., Rohlfing, K. J., & Raczaszek-Leonardi, J. (2016). *Vocal interactions at the dawn of communication: The emergence of mutuality and complementarity in mother-infant interaction*. Joint IEEE international conference on development and learning and epigenetic robotics ICDL-EpiRob (pp. 288–293).

Levinson, S. (1995). *Interactional biases in human thinking. Social Intelligence and Interaction: Expressions and implications of the social bias in human intelligence*. Cambridge University Press. https://doi.org/10.1017/CBO9780511621710.014

Levinson, S. (2006). Cognition at the heart of human interaction. *Discourse Studies*, *8*(1), 85–93. https://doi.org/10.1177/1461445606059557

Levinson, S. (2013). Action formation and ascription. In *The handbook of conversation analysis* (pp. 101–130). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118325001.ch6

Levinson, S. (2016). Turn-taking in human communication—Origins and implications for language processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Levinson, S., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*, 731. https://doi.org/10.3389/fpsyg.2015.00731

Lewedag, V., Oller, D., & Lynch, M. (1994). Infants' vocalization patterns across home and laboratory environments. *First Language*, *14*(42–43), 49–65. https://doi.org/10.1177/014272379401404204

Lewis, M., Mathur, M., VanderWeele, T., & Frank, M. C. (2020). The puzzling relationship between multi-lab replications and meta-analyses of the rest of the literature. *PsyArXiv*. https://doi.org/10.31234/osf.io/pbrdk

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

Lopez, L., Walle, E., Pretzer, G., & Warlaumont, A. (2020). Adult responses to infant prelinguistic vocalizations are associated with infant vocabulary: A home observation study. *PLoS One*, *15*(11), e0242232. https://doi.org/10.1371/journal.pone.0242232

Malin, D., Ludvig, I., Satu, K., Linnea, W., Helena, Z., Anna, E., & Ulrika, M. (2011). Parent-child interaction: Relationship between pause duration and infant vocabulary at 18 months. *Fonetik 2011*, *1*(1), 101–104. Proceedings from Fonetik 2011: Speech, music and hearing; quarterly progress and status report.

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

Marklund, U., Marklund, E., Lacerda, F., & Schwarz, I.-C. (2015). Pause and utterance duration in child-directed speech in relation to child vocabulary size. *Journal of Child Language*, *42*(5), 1158–1171. https://doi.org/10.1017/S0305000914000609

McFarland, D. H. (2001). Respiratory markers of conversational interaction. *Journal of Speech, Language, and Hearing Research*, *44*(1), 128–143. https://doi.org/10.1044/1092-4388(2001/012)

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.

Naigles, L. R., & Fein, D. (2017). Looking through their eyes: Tracking early language comprehension in ASD. In *Innovative investigations of language in autism spectrum disorder* (pp. 49–69). Walter de Gruyter GmbH. https://doi.org/10.1037/15964-004

Navarro, D. J. (2021). If mathematical psychology did not exist we might need to invent it: A comment on theory building in psychology. *Perspectives on Psychological Science*, *16*(4), 707–716. https://doi.org/10.1177/1745691620974769

Newman, L., & Smit, A. (1989). Some effects of variations in response-time latency on speech rate, interruptions, and fluency in children's speech. *Journal of Speech and Hearing Research*, *32*(3), 635–644. https://doi.org/10.1044/jshr.3203.635

Nikolaus, M., Maes, J., Auguste, J., Prévot, L., & Fourtassi, A. (2021). Large-scale study of speech acts' development using automatic labelling. *PsyArXiv*. https://doi.org/10.31234/osf.io/j4smd

Northrup, J. B., & Iverson, J. M. (2015). Vocal coordination during early parent-infant interactions predicts language outcome in infant siblings of children with autism spectrum disorder. *Infancy*, *20*(5), 523–547. https://doi.org/10.1111/infa.12090

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

Ochi, K., Ono, N., Owada, K., Kojima, M., Kuroda, M., Sagayama, S., & Yamasue, H. (2019). Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PLoS One*, *14*(12), e0225377. https://doi.org/10.1371/journal.pone.0225377

Ochs, E., Kremer-Sadlik, T., Sirota, K. G., & Solomon, O. (2004). Autism and the social world: An anthropological perspective. *Discourse Studies*, *6*(2), 147–183. https://doi.org/10.1177/1461445604041766

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, *372*, n71. https://doi.org/10.1136/bmj.n71

Pezzulo, G., Donnarumma, F., Dindo, H., D'Ausilio, A., Konvalinka, I., & Castelfranchi, C. (2019). The body talks: Sensorimotor communication and its brain and kinematic signatures. *Physics of Life Reviews*, *28*, 1–21. https://doi.org/10.1016/j.plrev.2018.06.014

Pickering, M., & Garrod, S. (2021). *Understanding dialogue*. Cambridge University Publisher. https://www.cambridge.org/us/academic/subjects/languages-linguistics/psycholinguistics-and-neurolinguistics/understanding-dialogue-language-use-and-social-interaction

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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*, *285*(1880), 20180598. https://doi.org/10.1098/rspb.2018.0598

Rast, P., & Ferrer, E. (2018). A mixed-effects location scale model for dyadic interactions. *Multivariate Behavioral Research*, *53*(5), 756–775. https://doi.org/10.1080/00273171.2018.1477577

Reissland, N., & Stephenson, T. (1999). Turn-taking in early vocal interaction: A comparison of premature and term infants' vocal interaction with their mothers. *Child: Care, Health and Development*, *25*(6), 447–456. https://doi.org/10.1046/j.1365-2214.1999.00109.x

Rich, P., de Haan, R., Wareham, T., & van Rooij, I. (2021). How hard is cognitive science? *PsyArXiv*. https://doi.org/10.31234/osf.io/k79nv

Riest, C., Jorschick, A. B., & de Ruiter, J. P. (2015). Anticipation in turn-taking: Mechanisms and information sources. *Frontiers in Psychology*, *6*, 89. https://doi.org/10.3389/fpsyg.2015.00089

Rochet-Capellan, A., & Fuchs, S. (2014). Take a breath and take the turn: How breathing meets turns in spontaneous dialogue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658), 20130399. https://doi.org/10.1098/rstb.2013.0399

Ruiter, J., Mitterer, H., & Enfield, N. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535. https://doi.org/10.1353/lan.2006.0130

Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. In J. Schenkein (Ed.), *Studies in the organization of conversational interaction* (pp. 7–55). Academic Press. https://doi.org/10.1016/B978-0-12-623550-0.50008-2

Savelkoul, E. M., Zebrowski, P. M., Feldstein, S., & Cole-Harding, S. (2007). Coordinated interpersonal timing in the conversations of children who stutter and their mothers and fathers. *Journal of Fluency Disorders*, *32*(1), 1–32. https://doi.org/10.1016/j.jfludis.2006.12.001

Sawyer, J., Matteson, C., Ou, H., & Nagase, T. (2017). The effects of parent-focused slow relaxed speech intervention on articulation rate, response time latency, and fluency in preschool children who stutter. *Journal of Speech, Language, and Hearing Research*, *60*(4), 794–809. https://doi.org/10.1044/2016_JSLHR-S-16-0002

Snow, C. E. (1977). The development of conversation between mothers and babies*. *Journal of Child Language*, *4*(1), 1–22. https://doi.org/10.1017/S0305000900000453

Stern, D., Beebe, B., Jaffe, J., & Bennett, S. (1977). The infant's stimulus world during social interaction: A study of caregiver behaviors with particular reference to repetition and timing. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 177–202). Academic Press.

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *British Medical Journal*, *338*, b2393. https://doi.org/10.1136/bmj.b2393

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., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology*, *3*(3), 289–299. https://doi.org/10.1080/17405620600879779

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), 20150370. 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., Fawcett, C., & Liszkowski, U. (2014). Infants' expectations about gestures and actions in third-party interactions. *Frontiers in Psychology*, *5*, 321. https://doi.org/10.3389/fpsyg.2014.00321

Todt, D., & Hultsch, H. (1999). How nightingales develop their vocal competence. *Proc. 22nd Int. Ornithological Congress* (pp. 193–215).

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press. https://doi.org/10.2307/j.ctvjsf4jc

Traag, V. A., & Bruggeman, J. (2009). Community detection in networks with positive and negative links. *Physical Review E*, *80*(3), 036115. https://doi.org/10.1103/PhysRevE.80.036115

Trecca, F., Tylén, K., Højen, A., & Christiansen, M. (2021). Danish as a window onto language processing and learning. *Language acquisition*, *71*(3), 799–833.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(1), 1–67. https://doi.org/10.18637/jss.v045.i03

Waade, P. T., Enevoldsen, K. C., Simonsen, A., & Fusaroli, R. (2020). Introducing tomsup: Theory of mind simulations using python. *PsyArXiv*. https://doi.org/10.31234/osf.io/zcgkv

Warlaumont, A., Oller, D., Dale, R., Richards, J. A., Gilkerson, J., & Xu, D. (2010). Vocal interaction dynamics of children with and without autism. *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 121–126).

Warlaumont, A., Richards, J., Gilkerson, J., & Oller, D. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023

Weed, E., Fusaroli, R., Tranbjerg, J., Fein, D., & Naigles, L. (2017). Conversational dynamics in a longitudinal corpus of caregiver-child interactions. Conference paper presented at International Meeting for Autism Research, San Francisco, California, USA.

Welkowitz, J., Bond, R., Feldman, L., & Tota, M. (1990). Conversational time patterns and mutual influence in parent-child interactions—A time-series approach. *Journal of Psycholinguistic Research*, *19*(4), 221–243. https://doi.org/10.1007/BF01077258

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, *12*(6), 957–968. https://doi.org/10.3758/bf03206432

Yaruss, J. (1997). Utterance timing and childhood stuttering. *Journal of Fluency Disorders*, *22*(4), 263–286. https://doi.org/10.1016/S0094-730X(97)00023-5

Zlochower, A., & Cohn, J. (1996). Vocal timing in face-to-face interaction of clinically depressed and nondepressed mothers and their 4-month-old infants. *Infant Behavior and Development*, *19*(3), 371–374. https://doi.org/10.1016/S0163-6383(96)90035-1

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** 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*, 93, 1181–1200. https://doi.org/10.1111/cdev.13754

14678624, 2022, 4, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13754 by Ku Leuven, Wiley Online Library on [19/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License