UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Development of Flexible Role-Taking in Conversations Across Preschool

Permalink
https://escholarship.org/uc/item/8gd7485s

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Peirolo, Morgane
Xu, Zihan
Fourtassi, Abdellah

Publication Date
2024

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Development of Flexible Role-Taking in Conversations Across Preschool

Morgane Peirolo (morgane.peirolo@gmail.com)
Zihan XU (zhxu98@gmail.com)
Abdellah Fourtassi (abdellah.fourtassi@gmail.com)

Aix-Marseille Université, Université de Toulon, CNRS, LIS, Marseille, France

Abstract

The paper investigates the development of conversational skills in preschool children, focusing on their ability to adopt flexible roles in dialogues. We specifically analyze children's coordinated behavior in question-response-follow-up sequences, both as Initiators and Responders, using a longitudinal French corpus of child-caregiver spontaneous interactions. While preschool children showed growing sophistication in their ability to initiate and respond appropriately within conversations, they still had qualitative differences with adults, especially as initiators, suggesting further development beyond preschool. The findings contribute to our understanding of how conversational skills develop in early childhood and the role these skills play in broader cognitive and social development.

Keywords: Language acquisition, social coordination, conversational development

Introduction

Conversation is a coordinated activity: Interlocutors are sensitive and responsive to each other's signals, facilitating shared understanding and, therefore, successful communication (H. H. Clark, 1996; Pickering & Garrod, 2021). How can we characterize the development of this skill? One way is to measure children's ability to *align* their verbal and nonverbal signals with the interlocutors (e.g., Chieng et al., 2024; Fusaroli et al., 2023; Mazzocconi et al., 2023; Misiek & Fourtassi, 2022; Misiek et al., 2020).

Nevertheless, alignment – which reflects the ability to mirror the interlocutor – is only one facet of coordination. Another key ingredient is learning to play *complementary* conversational roles when appropriate (e.g., Fusaroli et al., 2014; Schegloff, 1986). Take the following example:

- Parent: What are you doing? (Question)
- Child: I am drawing. (Response)
- Parent: Oh, nice! (Follow-up)

This sequence is initiated by the parent who asks a question. This question invites the child to respond, which they did. The parent then follows up on the child's response with evaluative feedback. In this example, the parent and child are successfully collaborating while not doing identical things; rather, they are fulfilling the complementary roles afforded by this conversational sequence. Had the child just mimicked the parent, responding to a question with another question, the coordination would not have been perceived as successful.

For children to become skillful conversational partners, they must develop the ability to play different conversational roles flexibly, depending on the context (Agrawal et al., 2024; Nikolaus et al., 2022; Snow et al., 1996). In particular, they should master the ability to frequently alternate between *initiating* and *responding* roles, essential for maintaining communicative flow and coordination.

Here is an example, similar in structure to the above, but where the roles are reversed:

- Child: Are we going to the park today? (Question)
- Parent: Yes (Response)
- Child: Yay! (Follow-up)

Related work

We aim to study children's coordination skills in terms of learning to play complementary conversational roles within a given conversational sequence and to be flexible in switching roles as initiators or responders. We study how these skills develop across preschool by analyzing a longitudinal French corpus of child-caregiver spontaneous interactions at home.

To make systematic comparisons within and across age, we focus on conversational sequences similar to the ones shown in the examples above, that is, sequences initiated by a question. Indeed, questions afford a structured sequence of back-and-forth: Question, Response, and Follow-up (hereafter QRF), allowing us to test children's collaborative behavior (or lack thereof) in different roles.

Children start both asking questions and answering them relatively early, starting from around two years of age (Casillas et al., 2016; Chouinard et al., 2007; Ervin-Tripp, 1979; Gallagher, 1981; Garvey & Hogan, 1973; Mueller, 1972; Snow et al., 1996). This makes QRF sequences ideal for investigating children's conversational development across the preschool period.

While our main goal is to use QRFs as a window into children's early coordination skills (Stivers et al., 2018), our study is relevant to the literature on children's social learning, where QRFs are hypothesized to be the locus of important knowledge acquisition mechanisms (e.g., Kurkul & Corriveau, 2018; Yu et al., 2019).

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Current study

The study's main novelty is its systematic, longitudinal comparison of child-initiated vs. caregiver-initiated QRFs across the preschool years. Previous work has provided insights into children's question-asking or question-answering; however, to test flexible role-taking in conversations, we need to observe the same behavior in both situations. Further, by observing French-learning children, our work contributes to ongoing efforts to test the generality of our current knowledge of children's early conversational skills (obtained mainly in North America) to other cultures (see Gauvain & Munroe, 2020).

Methods

Data

We used data from the "Paris Corpus" (Morgenstern & Parisse, 2012). The corpus is made of longitudinal recordings of French children spontaneously interacting with their caregivers at home. The participants were videotaped (by a researcher) once a month for about an hour over a developmental period ranging from 1 to 5 years of age (with slight variation in this range across children).

The corpus contains data from 6 children (4 males and 2 females). We kept both female children (Anae and Madeleine) and picked two males, Adrien and Theophile, as preliminary inspection showed their video recordings to be of higher quality (the multimodal context was important for our annotation). All children belonged to middle-class families with college-level educated caregivers (more background about the participants can be found on the corpus website).[1]

As the total number of recordings is prohibitively too large to fully annotate, we sampled six transcripts per child, capturing the beginning, middle, and end of the developmental range under study. We ended up with 24 transcripts and their video recordings, which amounted to around 24 hours.

Figure 1 shows the number of dialog turns of interlocutors per transcript. Overall, the number of turns by the caregiver and the child was roughly similar. They were generally above 500 turns per transcript/hour even at the younger end of the developmental range.

Coding scheme

We took as a starting point the literature on child-initiated QRF sequences, especially the work by Kurkul and Corriveau (2018), who introduced a rather comprehensive coding scheme for QRF sequences in the context of early child-caregiver interaction. We adapted this code to be applied to both child-initiated and parent-initiated QRF.

Questions were categorized into three major groups: fact-based, explanation-based (both are information-seeking), and other, non-information-seeking questions (e.g., requests). The categories for Responses can be classified broadly as distinguishing relevant/contingent answers vs. non-contingent answers. Each was further subdivided to distinguish, for example, whether the answer was short, elaborate, or non-verbal. Finally, the Follow-up categories can also be broadly classified into contingent vs. non-contingent (relative to the question being asked), with further subdivision marking types such as evaluative feedback or follow-up questions. The full list of categories and sub-categories we used will be provided in the Appendix.

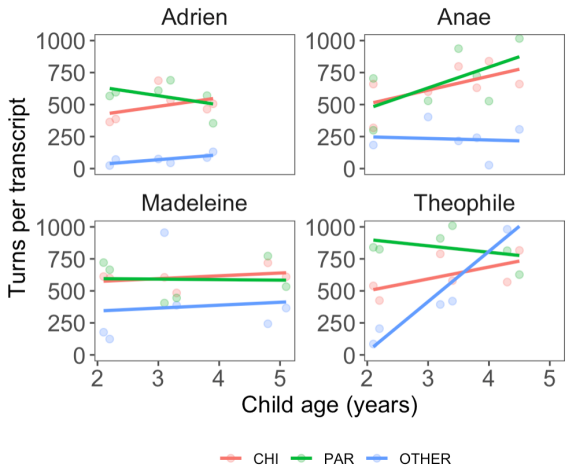

Figure 1: Total number of turns by children, caregivers, or other participants in the dialog (e.g., experimenter) per recording/transcript as a function of the child's age. Lines represent best linear fit.

Annotation Method and Reliability

Annotation Platform We facilitated the annotation process by designing a computer platform that parsed the transcripts turn by turn.[2] In the platform, each turn was accompanied by several drop-down lists made of QRF categories and sub-categories that the annotators could choose from. The platform offered multiple duplicate lists in case the same utterance could be classified into multiple categories (e.g., a question can be both a follow-up in the current QRF and an initiation in the next QRF). A screenshot of the platform will be provided in the Appendix.

Procedure Two human annotators did the annotation. The first rounds were dedicated to fine-tuning the annotation scheme. The transcripts used for this initial round were not included in the analysis. After the coding scheme was finalized, the annotation proceeded as follows. First, a random sample of 20% of the data was independently annotated to calculate inter-annotation agreement (see subsection below). After this step, and to optimize the annotation size, each annotator coded separate transcripts.

Each transcript was annotated twice: Once for child-initiated QRF and once for caregiver-initiated QRF. The an-

[1] The corpus link: https://phonbank.talkbank.org/access/French/Paris.html

[2] The turns were already segmented by the original authors providing the data, i.e., Morgenstern and Parisse (2012)

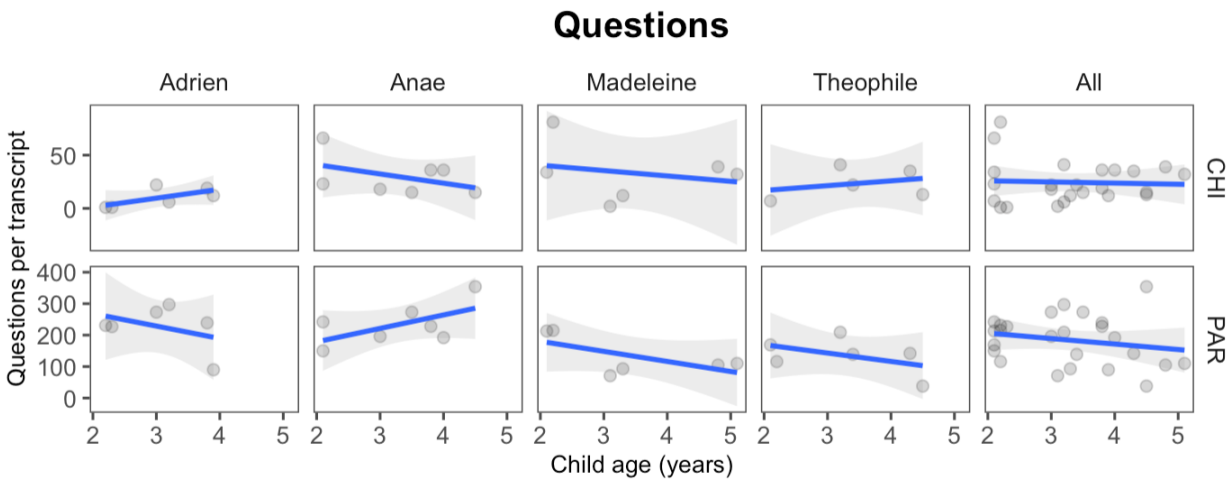

Figure 2: Total number of questions asked by children to caregivers and vice versa, as a function of the child's age. Dots represent this number per transcript, lines represent best linear fit, and grey bands represent their % 95 confidence intervals.

notation relied on the transcript and the corresponding videotapes as many situations required the multimodal context (e.g., the child nodding their head as a non-verbal response).

Inter-rater reliability Inter-annotation agreement was estimated based on a random sample of about 20% of the data, coded independently by each annotator. Cohen's kappa values were 0.8 for child-initiated QRF data and 0.9 for caregiver-initiated QRF data. The differences/discrepancies in this subset of data were resolved through discussion.

Analysis

The section studies the properties of questions, responses, and follow-ups when children are both in the initiator and responder roles.

Questions

Questions' Number Figure 2 shows the number of questions asked by children to parents (and vice versa). For children, the number of questions asked to parents varied generally between 0 and 50 per hour. The average number was 24.5 questions/hour. For the caregivers, the number of questions asked to children varied roughly between 100 and 300, with an average of 184.5 questions/hour.

The caregivers asked almost an order of magnitude more questions to children than the other way around – despite the fact that both children and caregivers contributed to the dialog with a roughly similar number of turns (Figure 1). This discrepancy shows that most of the QRF dynamics were initiated by caregivers. We found no evidence of change in this pattern across the preschool period.

Questions' types Figure 4 shows the question's types. The majority of questions asked by children were information-seeking (around 85% on average), including both fact-seeking questions (e.g., what-, who-, and where-questions) and explanation-seeking questions (e.g., why- and how- questions). The remaining (i.e., 15%) were non-information-seeking, such as requests for permission, help, or attention. We found, overall, a similar pattern in parents' questions.

Responses

Response contingency Figure 3 shows the proportion of *contingent* responses provided by children and adults. Contingency determines whether the response was coordinated and provided a relevant answer. We counted as non-contingent cases where the response was irrelevant or absent.

For all children, response contingency increases across preschool, approaching around 75 % average contingent responses by 4 to 5 years old. This observation was corroborated by statistical testing. We fit a mixed-effects logistic regression predicting whether a response is contingent ($= 1$) or non-contingent ($= 0$) as a function of child age.[3] We found `Age` to be a highly significant predictor $\beta = 0.43$ ($SE = 0.11$, $p < 0.001$).

As for responses provided by parents, data was much more noisy/variable (in part, due to the fact that children asked fewer questions per transcript, naturally leading to fewer overall responses by parents). That said, caregivers' contingency was generally lower than what we observed in children, it was on average generally around 50% per question. Unlike children's response contingency, there was no evidence that caregivers' behavior changed with development.

Response types We classified the (contingent) response sub-categories into three broad classes: short, elaborate, and non-verbal. The distribution is shown in Figure 5. We found most responses to be short and only a minority to be elaborate (in both children and parents). This finding mirrors the fact

[3]The model was specified as follows:
`Contingent_or_not ~ Age + (Age | child)`

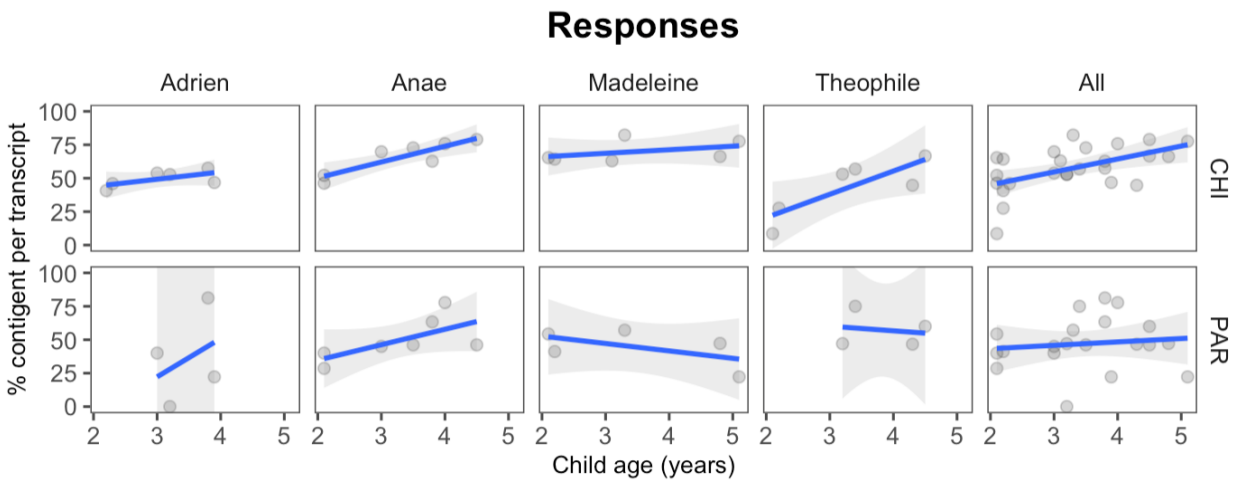

Figure 3: Percent of contingent responses made by children (top row) and by caregivers (bottom row), as a function of the child's age. Dots represent the percent per transcript, lines represent best linear fit, and grey bands represent their % 95 confidence intervals. For caregivers' responses (bottom row), we removed transcripts with only one response (which happens if the child only asked one question in the whole transcript) because this leads to spurious percentages (100% or 0%).

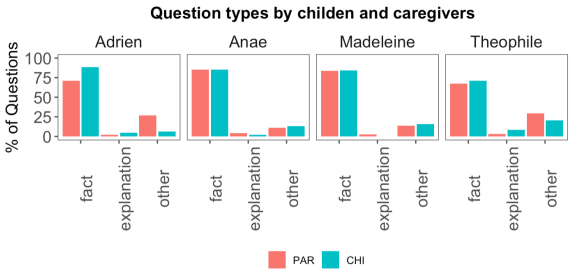

Figure 4: Percentage of question types for each child and caregiver across all transcripts. Types are fact-seeking questions (e.g., what-, who-, and where-questions), explanation-seeking (e.g., why- and how-questions), and other (e.g., request for permission and help).

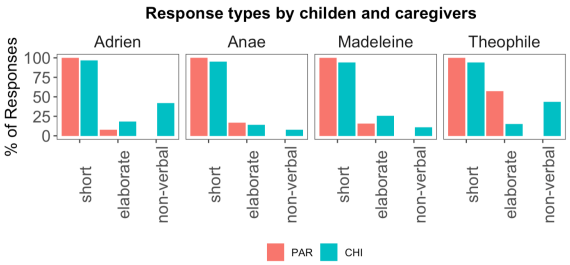

Figure 5: Percentage of response types for each child and caregiver across all transcripts. Types are classified here as short, elaborate, or non-verbal.

that most questions were fact-seeking (requiring rather short responses). A non-negligible proportion of children's contingent responses was non-verbal (this was more frequent in the early years, data not shown).

Follow-ups

Follow-up contingency Figure 6 shows the proportion of *contingent* follow-ups provided by children and adults. Here, contingency is measured relative to the question, reflecting the questioner's appreciation of the response. We counted as non-contingent cases where the follow-up was irrelevant to the question or absent. Children generally had a quite high rate of contingent follow-ups. The average is 57%. We observe little change across development. Parents also provided mostly contingent follow-ups (around 75% on overage).

Follow-up types The (contingent) follow-up subcategories could be classified into three broad classes: follow-up questions, providing one's own response to the question or/and adding more information to the interlocutor's response, and evaluation of the interlocutor's response (e.g., agreement or disagreement). The distribution of these three types is shown in Figure 7.

Unlike the case of question types (and response types to a lesser degree) where caregivers and children had similar distributions, here we found two main differences: a) caregivers followed up more with questions than children did, and b) children followed up with more own-responses than adults did. These two patterns were consistent across all dyads.

Interaction between Response and Follow-up

Above, we quantified contingency in response and follow-up relative to the question asked. Here, we examine the interaction between responses and follow-ups; that is, we examine

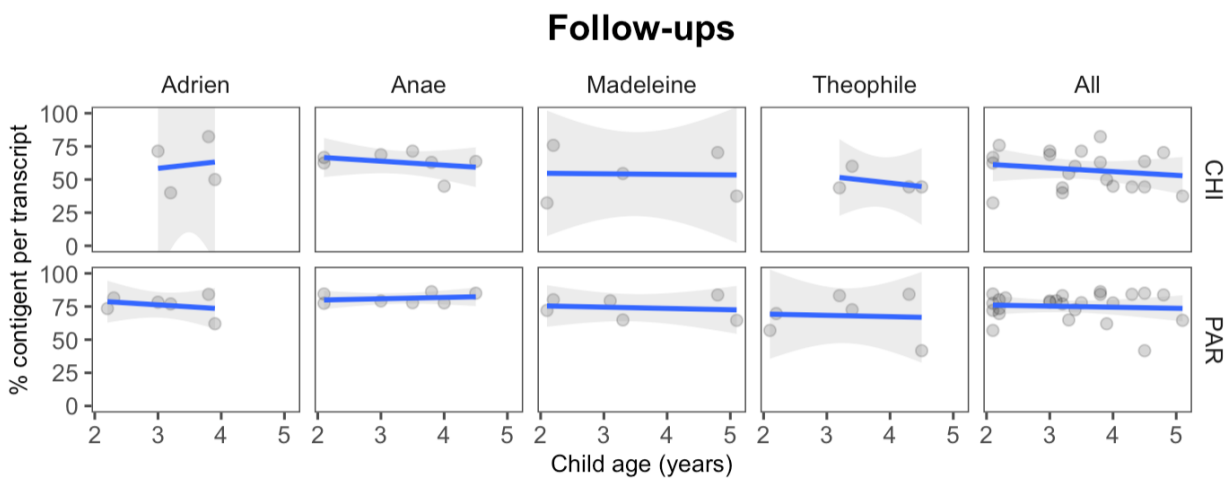

Figure 6: Percent of contingent follow-ups made by children (top row) and by caregivers (bottom row), as a function of the child's age. Dots represent the percent per transcript, lines represent best linear fit, and grey bands represent their % 95 confidence intervals. For children's responses (top row), we removed transcripts with only one follow-up (which happens if the child only asked one question in the whole transcript), as this leads to spurious percentages (100% or 0%).

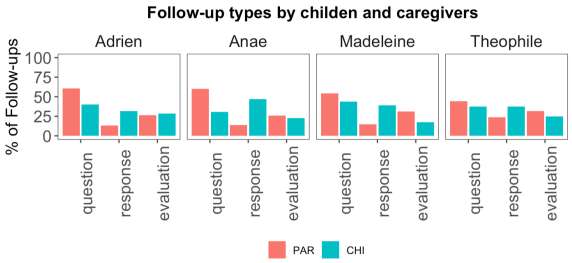

Figure 7: Percentage of follow-up types for each child and caregiver across all transcripts. Types are: follow-up questions, own response to the question asked, and evaluation of the interlocutor's response (e.g., agreement or disagreement).

how caregivers' contingency (or lack thereof) influences the child's follow-up (and vice-versa). To this end, we compared follow-ups after contingent vs. non-contingent responses. The results are shown in Figure 8.

For children, we found that *more* contingent follow-ups were given following *non-contingent* responses from caregivers. To verify this observation statistically, we fit a mixed-effects logistic regression predicting whether a follow-up by children was contingent ($= 1$) or not ($= 0$) as a function of response contingency (also binary) and Age.[4] Indeed, we found a *negative* effect of `Response_contingency`: $\beta = -1.70$ ($SE = 0.72$, $p = 0.02$). The interaction term `Response_contingency * Age` was not significant: $\beta = 0.27$ ($SE = 0.21$, $p = 0.20$), meaning there was no evidence of change across development. Figure 8 (top row) shows that this effect was consistent within all children except for Adrien.

For adults, we found – interestingly – the opposite pattern: More contingent follow-ups were given following *contingent* responses from children. To verify this observation statistically, we fit a similar mixed-effects logistic regression as we did with children. Confirming our qualitative observations, we found a *positive* effect of `Response_contingency`: $\beta = 0.82$ ($SE = 0.34$, $p = 0.02$). The interaction term `Response_contingency * Age` was not significant in this case, either: $\beta = -0.13$ ($SE = 0.10$, $p = 0.19$), meaning there was no evidence of change across development in this case, either. Figure 8 (bottom row) shows that this effect was consistent within all caregivers (with no exception).

Discussion

A crucial aspect of children's linguistic and social cognitive development is learning to hold *coordinated* conversations with people around them. Part of this coordination involves the child and interlocutor doing similar things (e.g., mirroring each other's choice of words), but another part requires playing complementary roles. This paper investigated the latter. Specifically, we examined how preschool children behave in QRF sequences initiated by themselves or the caregiver, testing their ability to coordinate in different roles. Analysis of a French corpus of spontaneous child-caregiver interaction led to the following findings.

Questions and Responses in child-caregiver conversations QRFs are commonplace, structuring a large chunk of child-caregiver interactions (question-asking for parents and question-answering for children, see Figure 1 and 2). This provides a posteriori validation for using QRFs as a general window into children's emergent conversational skills (see

[4]The model was specified as follows:
`Followup_contingency ~ Response_contingency*Age + (1 | child)`

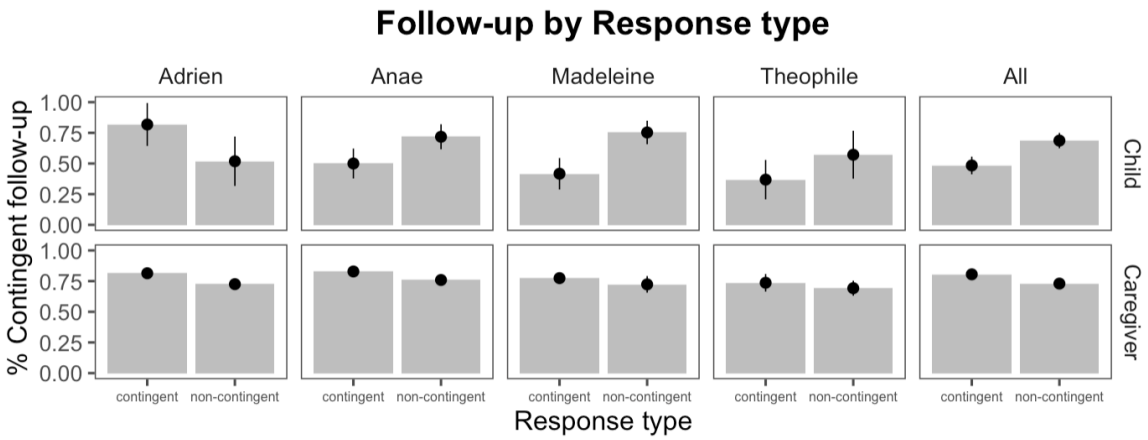

Figure 8: Percent of contingent follow-ups per response contingency status (summed over all ages and transcripts for maximal statical power). Dots and ranges represent the mean and % 95 confidence intervals.

also Stivers et al., 2018)

Responses were our first test of children's coordination in a complementary role. We found that most children gave well-coordinated, contingent responses more than 50% of the time, even at the youngest age (around two years old), confirming results obtained with English-learning children of similar ages (Dore, 1977; Gallagher, 1981). Further, we found that all children improved in contingency across preschool, corroborating evidence from previous studies comparing younger and older preschool children (e.g., Garvey, 1975).

Evidence for flexible role-taking In addition to responses, testing children's follow-ups was crucial, as this allowed us to observe children playing different conversational roles in the same conversations. Their follow-ups were highly contingent, starting from the earliest ages. This finding – together with the high response contingency rate – indicates precocious sophisticated coordination skills, allowing children to adapt flexibly to different conversational roles, namely the roles of a responder and of an initiator.

The first (i.e., the responder) requires attention to the interlocutor's intents and willingness to play along, largely on the interlocutor's terms. Indeed, it is the question that elicits the response. Further, the question constrains responses both in form (e.g., a yes-no question requires a different response format than, say, a wh-question) and in content (i.e., the response must address the question). In contrast, the second role (i.e., the initiator) requires the ability to evaluate the interlocutor's responses to one's own question and to be proactive in signaling – via follow-ups – (dis-)satisfaction with it and, if necessary, willingness to pursue the interaction further to reach a better communicative outcome.

Both the initiator and responder roles are crucial for communicative coordination, and interlocutors often switch back and forth between these roles in the course of a single conversation. The results reported in this work provide evidence of this role-taking flexibility at a young age, improving across preschool.

Room for Development? There were differences between children and adults, especially in the initiator's role. First, children initiated fewer questions than parents did. In fact, French children in our corpus asked even fewer questions (25 per hour) than previously reported for English-speaking children in a generally similar interactive context (108/hour, Chouinard et al., 2007). This difference, however, can be due to the properties of the specific context of the recorded interactions (Bodur et al., 2023; Dideriksen et al., 2023; Jiang et al., 2022; Wynn et al., 2024). Second, children and caregivers showed divergent patterns both in terms of follow-up types (e.g., caregivers asked more follow-up questions than children did) and, more interestingly, in terms of how follow-ups were prompted by the interlocutor's responses (Figure 8). While parents showed a relatively high rate of follow-ups, whether or not the child's response was contingent (with a slight increase for the contingent ones), children, in contrast, did not always follow up after contingent responses and showed more evidence for *persistence* following non-contingent, unsatisfactory responses. Forms of persistence have been reported before, not only in early conversations (Chouinard et al., 2007; Frazier et al., 2009; Kurkul & Corriveau, 2018) but also in non-verbal infant-parent interactions (Tronick et al., 1978) and in other species (Dingemanse & Enfield, 2024). Its early emergence may have to do with its developmentally urgent role – at least from the point of view of the initiator – in repairing ongoing communication. Children later learn social norms that encourage follow-up (e.g., showing appreciation) even when one's questions have been satisfactorily attended to.

One important future research direction is to investigate the implications of flexible role-taking in early interactions (or lack thereof) on the effectiveness of social learning (E. V. Clark, 2022; Kurkul & Corriveau, 2018; Nikolaus & Fourtassi, 2021, 2023; Yu et al., 2019).

Acknowledgement

This work was carried out within the Institute of Convergence ILCB (ANR-16-CONV-0002). It has benefited from support from the French government (France2030), managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix Marseille University (A*MIDEX). Furthermore, this study was also supported by the ANR MACOMIC (ANR-21-CE28-0005-01) grant. Finally, we thank Charlie Hallart for helping with the annotation and the refinement of the coding scheme.

References

Agrawal, A., Favre, B., & Fourtassi, A. (2024). Communicative intent coordination in child-caregiver interactions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*.

Bodur, K., Nikolaus, M., Prévot, L., & Fourtassi, A. (2023). Using video calls to study children's conversational development: The case of backchannel signaling. *Frontiers in Computer Science*, *5*.

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.

Chieng, A. C. J., Wynn, C. J., Wong, T. P., Barrett, T. S., & Borrie, S. A. (2024). Lexical alignment is pervasive across contexts in non-weird adult–child interactions. *Cognitive Science*, *48*(3), e13417.

Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, i–129.

Clark, E. V. (2022). Language is acquired in interaction. In *Algebraic structures in natural language* (pp. 77–94). CRC Press.

Clark, H. H. (1996). *Using language*. Cambridge university press.

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2023). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General*, *152*(3), 864.

Dingemanse, M., & Enfield, N. (2024). Interactive repair and the foundations of language. *Trends in Cognitive Sciences*, *28*(1), 30–42.

Dore, J. (1977). "oh them sheriff": A pragmatic analysis of children's responses to questions. In *Child discourse* (pp. 139–163). Elsevier.

Ervin-Tripp, S. (1979). Children's verbal turn-taking. *Developmental pragmatics*, 391–414.

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult–child conversation. *Child development*, *80*(6), 1592–1611.

Fusaroli, R., Raczaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157.

Fusaroli, R., Weed, E., Rocca, R., Fein, D., & Naigles, L. (2023). Repeat after me? both children with and without autism commonly align their language with that of their caregivers. *Cognitive Science*, *47*(11), e13369.

Gallagher, T. M. (1981). Contingent query sequences within adult–child discourse. *Journal of Child Language*, *8*(1), 51–62.

Garvey, C. (1975). Requests and responses in children's speech. *Journal of child language*, *2*(1), 41–63.

Garvey, C., & Hogan, R. (1973). Social speech and social interaction: Egocentrism revisited. *Child Development*, 562–568.

Gauvain, M., & Munroe, R. L. (2020). 10 children's questions in social and cultural perspective. *The questioning child: Insights from psychology and education*, *183*.

Jiang, H., Frank, M. C., Kulkarni, V., & Fourtassi, A. (2022). Exploring patterns of stability and change in caregivers' word usage across early childhood. *Cognitive Science*, *46*(7), e13177.

Kurkul, K. E., & Corriveau, K. H. (2018). Question, explanation, follow-up: A mechanism for learning from others? *Child Development*, *89*(1), 280–294.

Mazzocconi, C., O'Brien, B., El Haddad, K., Bodur, K., & Fourtassi, A. (2023). Differences between mimicking and non-mimicking laughter in child-caregiver conversation: A distributional and acoustic analysis. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*(45).

Misiek, T., Favre, B., & Fourtassi, A. (2020, November). Development of multi-level linguistic alignment in child-adult conversations. In E. Chersoni, C. Jacobs, Y. Oseki, L. Prévot, & E. Santus (Eds.), *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 54–58). Association for Computational Linguistics.

Misiek, T., & Fourtassi, A. (2022). Caregivers exaggerate their lexical alignment to young children across several cultures. *Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue - Full Papers*.

Morgenstern, A., & Parisse, C. (2012). The paris corpus. *Journal of French language studies*, *22*(1), 7–12.

Mueller, E. (1972). The maintenance of verbal exchanges between young children. *Child Development*, 930–938.

Nikolaus, M., & Fourtassi, A. (2021). Modeling the interaction between perception-based and production-based learning in children's early acquisition of semantic knowledge. *Proceedings of the 25th Conference on Computational Natural Language Learning*.

Nikolaus, M., & Fourtassi, A. (2023). Communicative feedback in language acquisition. *New Ideas in Psychology*, *68*, 100985.

Nikolaus, M., Maes, E., Auguste, J., Prévot, L., & Fourtassi, A. (2022). Large-scale study of speech acts' development in early childhood. *Language Development Research*, *2*(1), 268–304.

Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press.

Schegloff, E. A. (1986). The routine as achievement. *Human studies*, *9*(2-3), 111–151.

Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning how to say what one means: A longitudinal study of children's speech act use. *Social Development*, *5*(1), 56–84.

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

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child psychiatry*, *17*(1), 1–13.

Wynn, C. J., Barrett, T. S., & Borrie, S. A. (2024). Conversational speech behaviors are context dependent. *Journal of Speech, Language, and Hearing Research*.

Yu, Y., Bonawitz, E., & Shafto, P. (2019). Pedagogical questions in parent–child conversations. *Child development*, *90*(1), 147–161.

A - Annotation scheme

Child-initiated QRF coding scheme

CHI-Initiation:Information-seeking:fact-based
CHI-Initiation:Information-seeking:explanation-seeking
CHI-Initiation:Non-information-seeking

PAR-Response:Answer:no-explanation-needed
PAR-Response:Answer:with-explanation-low-quality
PAR-Response:Answer:with-explanation-high-quality
PAR-Response:Non-answer:declarative
PAR-Response:Non-answer:interrogative
PAR-Response:No-response

CHI-Follow-up:Signal-agreement
CHI-Follow-up:Signal-disagreement
CHI-Follow-up:Repeat-answer
CHI-Follow-up:Followup-question
CHI-Follow-up:Re-ask-initial-question
CHI-Follow-up:Own-response
CHI-Follow-up:Add-info
CHI-Follow-up:No-followup
CHI-Follow-up:Non-followup
CHI-Follow-up:No-label

Parent-initiated QRF coding scheme

PAR-Initiation:Information-seeking:fact-based
PAR-Initiation:Information-seeking:explanation-based
PAR-Initiation:Non-Information-seeking

CHI-Response:Answer:no-explanation-needed
CHI-Response:Answer:with-explanation-low-quality
CHI-Response:Answer:with-explanation-high-quality
CHI-Response:Answer:non-verbal
CHI-Response:Non-answer:declarative
CHI-Response:Non-answer:interrogative
CHI-Response:No-response
CHI-Response:No-label

PAR-Follow-up:Signal-agreement
PAR-Follow-up:Signal-disagreement
PAR-Follow-up:Repeat-answer
PAR-Follow-up:Followup-question
PAR-Follow-up:Re-ask-initial-question
PAR-Follow-up:Give-response
PAR-Follow-up:Add-info
PAR-Follow-up:No-followup
PAR-Follow-up:Non-followup

Figure 9: The coding schemes for both children- and parent-initiated QRF sequences.

B - Annotation platform

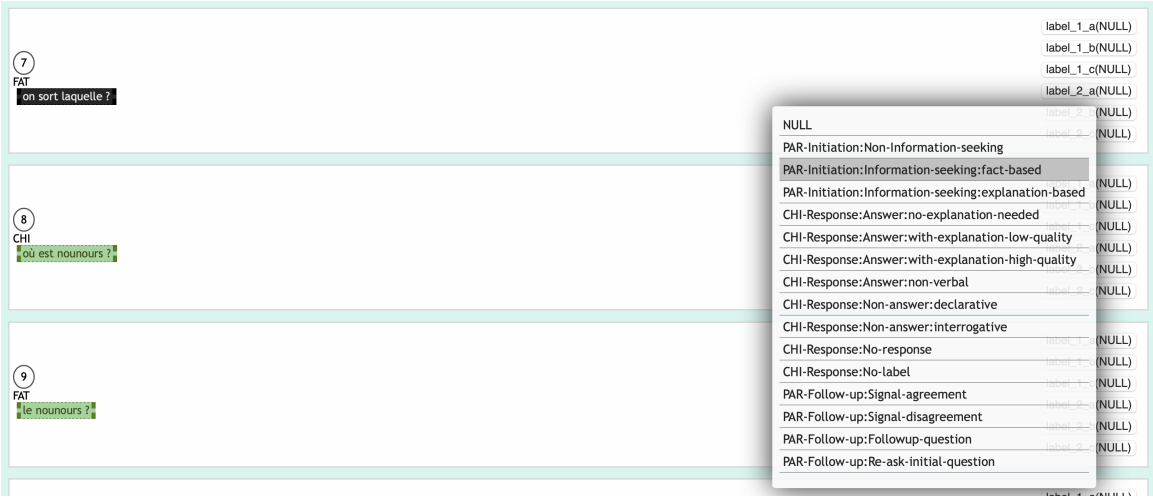

Figure 10: A screenshot of the online annotation platform. Given a dialog transcript, the platform parses it turn by turn and offers drop-down menus of QRF categories form which the annotator can choose.