Turn-timing in naturalistic mother-child interactions: A longitudinal perspective

Marisa Casillas[a,c], Elma Hilbrink[a,d], Susan C. Bobb[b], Eve V. Clark[c], Merideth Gattis[d], & Stephen C. Levinson[a]
a. MPI Nijmegen for Psycholinguistics,
b. Northwestern University, c. Stanford University, d. Cardiff University
Corresponding email: `Marisa.Casillas@mpi.nl`

Abstract

Combining data from two longitudinal studies of young children, we track the development of turn-timing in spontaneous infant-caregiver interactions. We focus on three aspects of timing: overlap, gap, and delay marking. We find evidence for early development of turn-timing skills, in-line with the Interaction Engine Hypothesis.

Part and parcel of learning a language is learning how to *use* it. Conversation is our first and primary mode of language use, and determines the form of children's linguistic input. But participating in conversation is not trivial; it requires interactants to weave together linguistic, non-verbal, and interactional information in real time, both while speaking and listening. Places of turn transition—when one speaker stops and the next one can start—are especially difficult since the interactants must coordinate on who will speak next, and when. Nonetheless, adults manage to take turns with apparent ease; their turn-transitions occur with minimal vocal overlap and gap between spoken turns. When speakers can't respond with immediate timing, they often delay their turn with markers such as *uh* and *um*. These patterns of turn-timing have been demonstrated in cultures around the globe, and thus appear to undergird human conversation (Clark & Fox Tree, 2002; Stivers et al., 2009). Cross-cultural universals in interactive structure are predicted by the Interaction Engine hypothesis, which suggests that human interactive abilities developed earlier and independently from linguistic abilities (IEH; Levinson, 2006). Applied to infant development, this same hypothesis predicts that infants begin to master interactive skills early and independently from their linguistic skills. We tested this idea by analyzing turn-timing in spontaneous interactions between English-speaking mothers and their children from 0;3 to 3;4. In-line with the IEH, we found that three aspects of turn-timing—vocal overlap avoidance, silent gap minimization, and marking response delays—emerge early in development and interact with children's linguistic planning once they begin to speak.

We analyzed turn-timing in two longitudinal free play corpora: (C1) 10-min in-lab recordings for 12 infants at 0;3, 0;4, 0;5, 1;0, and 1;6 (Ellis-Davies et al., 2012), and (C2) 1-hour at-home recordings for 5 children at 1;8, 2;0, 2;4, 2;8, 3;0 and 3;4 (Demuth, Culbertson, & Alter, 2006). In the first corpus we measured the timing of all transitions between vocalizations by the mother and baby.[1] In the second corpus, we measured the timing between 30 questions and answers for each child at each time point, and further coded each response for its complexity and markers of delay. We also measured the silent gap following turn-initial delay markers and preceding the rest of the turn (e.g., the '..' in "um .. that one"). As is typical of infant-parent interaction (Henning, Striano, & Lieven, 2005), most of the transitions from mother to baby in the first corpus were formatted as questions or ended in tag questions. Thus, the data from both corpora primarily represent turn-timing behavior in question-answer pairs.

Children and mothers took turns vocalizing throughout our sample (0;3–3;4). But, before 0;5, children frequently came in too early; they overlapped their vocalizations with the end of their mothers' nearly 40% of the time. At 0;5, children's overlaps began to decrease, matching the mothers by 1;6, and falling below them, to approximately 4%, by the first sample of the second corpus at 1;8 (Figure 1). This may suggest that children begin to avoid overlap at 0;5, respecting the norm of "one speaker at a time" (Sacks, Schegloff, & Jefferson, 1974).

[1] Except transitions from mother to baby when the baby's turn constituted a burp, sneeze, cough, etc.

Because of children's frequent overlap early on, their average turn-timing appears almost adult-like. Quick turn-timing in the first five months has also been reported in prior work (e.g., Ginsburg & Kilbourne, 1988), however these results are likely due to children's high frequency of overlapped starts during this early period. If we instead look at children's *gaps* (non-overlapped starts) with time, we see a clearer picture. Children start out on par with their mothers, but show significantly longer gaps at the 12-month sample. This increase in gap duration slowly tapers off over the rest of the sample until children converge with their mothers' timing again at 2;8 (Figure 2). The non-linear trajectory of gap timing (i.e., rise-then-fall) peaks near the onset of children's first words. If children's slower timing were really due to linguistic planning, we should find that more complex responses have longer gaps than less complex ones. We confirmed this with a linear mixed effects model of turn-timing in the second corpus, finding that more complex answers yielded longer gap durations (yes-no vs. wh-, single nominal vs. inflected phrase, $p<.001$) for children's, but not mothers' answers. This suggests that children may begin to minimize their gaps in the first year, but that the onset of speech may create significant planning costs and disrupts their ability to give an immediate response. Because of this, it may become crucial for children to mark their delays in speaking after 1;0.

Turn-initial delay markers (e.g., *uh*, *um*, prolongation, and repetition) emerged by 2;0 for all five children in the second corpus. Turns beginning with delay markers had significantly more linguistic material than those without, suggesting that children used delay markers when planning more complex responses ($p<.01$). Delay-marked responses were more complex, and so should have shown slower timing overall. However, the delay markers acted to buffer children's extra planning costs effectively, so that turns beginning with delay markers were not significantly longer than those without. Finally, by 3;6, children began to mark delay differentially, just like adults do, using *um* for longer delays and more difficult planning compared to *uh* (Clark & FoxTree, 2002, see also Hudson Kam and Edwards, 2008; Table 1). Children's delay marking suggests that they are cued in to the temporal sensitivity of transferring the floor from one speaker to the next.

In sum, we find that three aspects of turn-timing—overlap, gap, and delay marking—emerge early in children's development. Overlap is acquired first, with children making strides toward adult-like overlap patterns at 0;5. Gaps, too, appear short in the first year, but the onset of speech may cause children to slow down before they improve their overall timing. After the onset of speech, children begin to mark delay, holding the floor when planning complex responses and indicating their attention to the sensitivity of turn transitions. In-line with the IEH, the longitudinal outlook from these data support the idea that turn-timing skills develop early and independently from language, but also are consistent with the fact that, once children begin to speak, the linguistic and interactional systems must converge for children to continue developing adult-like conversational behavior.

References

H. H. Clark and J. E. Fox Tree. Using *uh* and *um* in spontaneous speaking. *Cognition*, 84(1):73–111, 2002.

K. Demuth, J. Culbertson, and J. Alter. Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech*, 49(2):137–173, 2006.

K. Ellis-Davies, E. Sakkalou, N. C. Fowler, E. E. Hilbrink, and M. Gattis. CUE: The continuous unified electronic diary method. *Behavior Research Methods*, 44(4), 1063–1078, 2012.

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

A. Henning, T. Striano, and E. V. Lieven. Maternal speech to infants at 1 and 3 months of age. *Infant Behavior and Development*, 28, 519–536, 2005.

C. L. Hudson Kam and N. A. Edwards. The use of uh and um by 3-and 4-year-old native English-speaking children: Not quite right but not completely wrong. *First Language*, 28(3):313–327, 2008.

S. C. Levinson. On the human 'interaction engine'. In N. J. Enfield and S. C. Levinson, editors, *Roots of human sociality: Culture, cognition and interaction*, pages 279–298. Berg, Oxford, England, 2006.

H. Sacks, E. A. Schegloff, and G. Jefferson. A simplest systematic for the organization of turn-taking for conversation. *Language*, 50:696–735, 1974.

T. Stivers, N. J. Enfield, P. Brown, ... and S.C. Levinson. Universals and cultural variation in turn-taking in conversation. *PNAS*, 106:10587–10592, 2009.

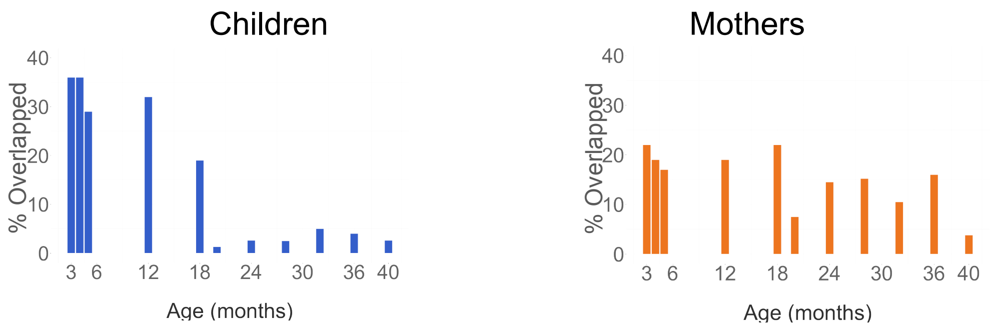

Figure 1. Percent overlapped vocalizations by children (left) and mothers (right) by children's age

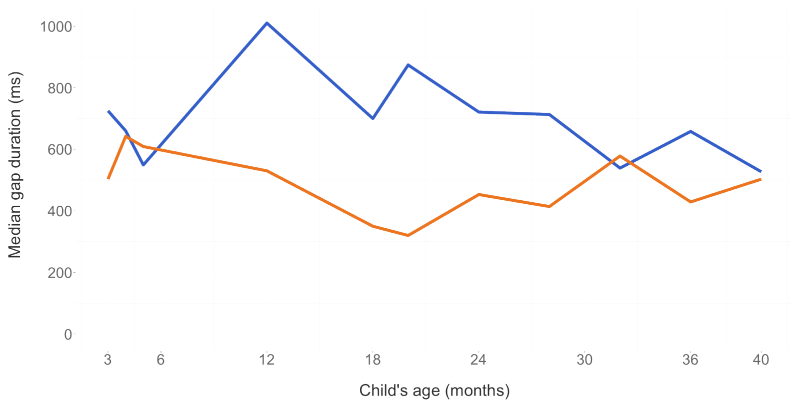

Figure 2. Median gap duration for children (top) and mothers (bottom) by children's age.

Delay marker	Pause duration	Speech fluency
uh	290 ms	73% fluent
um	450 ms	55% fluent

Table 1. Pause duration and speech fluency following children's use of *uh* and *um*.