Parents scaffold the formation of conversational pacts with their children

Ashley Leung
ashleyleung@uchicago.edu
Department of Psychology
University of Chicago

Robert D. Hawkins
rdhawkins@princeton.edu
Department of Psychology
Princeton University

Daniel Yurovsky
yurovsky@cmu.edu
Department of Psychology
Carnegie Mellon University
University of Chicago

Abstract

Adults readily form pacts, or temporary agreements about referent names, over the course of conversation. Young children fail to do so with peers, but recent evidence suggests that explicit feedback from adults may improve their performance (Matthews, Lieven, & Tomasello, 2007). Do parents naturally provide such structure in their conversations with children? Using a director-matcher paradigm, we first show that parents and children (ages 4, 6, 8) converge on increasingly accurate and efficient conversational pacts. Further, parents of younger children provide more interactive feedback. Finally, we analyze asymmetries in parents' and children's contributions, finding that pacts tend to originate with the parent, but are simplified by younger children. Together, these results support the idea that parents sensitively adapt their language to their children's developmental level to scaffold successful communication.

Keywords: communication; language development; parent-child interaction

Common ground is a foundational requirement for successful communication: Interlocutors must both understand each other and know that they understand each other (Bruner, 1985). As a consequence, even simple referential expressions are constructed collaboratively. When the meaning of an utterance is unclear, interlocutors will engage in negotiation, arriving at a *conversational pact* about how to think and talk about the intended referent (Clark & Wilkes-Gibbs, 1986). Adults fluently form, revise, and track the formation of pacts. For instance, Brennan and Clark (1996) showed that adults will use the appropriate level of specificity for a referent (e.g. "loafer" rather than "shoe" when another shoe is present), carry this forward into a new context with the same conversational partner even when the initial competitor is removed, then revert to "shoe" when a new partner is introduced.

Children, on the other hand, struggle substantially both to establish and to deploy conversational pacts. In a series of studies, Krauss and Glucksberg (1977) asked pairs of age-matched children to play a director-matcher game that required them to refer to novel objects (see also Krauss & Glucksberg, 1969). Across repeated interactions with the same novel objects, 5- and 6-year-old children showed very little evidence of pact formation: The matcher's accuracy in selecting the correct referent remained low, and the director did not reduce of the length of their referential expressions as adults do. Older children showed some improvement on both measures, but even 10- and 11-year-olds were less successful than a comparison group of adults.

While children struggle as directors, their performance can improve significantly when an adult offers explicit guidance on whether their referential expressions are ambiguous or otherwise insufficiently informative. For example, Deutsch & Pechmann (1982) showed that clarification questions could be used by adults to eventually elicit unambiguous descriptions from 6- and 9-year-olds and, to a lesser extent, from 3-year-olds. A recent training study by Matthews et al. (2007) replicated this finding in 2- to 4-year-olds, and further showed transfer to a different communication task. While these studies show how adult input may help children produce informative referential expressions, it remains unclear whether, or how, adults provide such feedback in the course of natural adult-child conversations.

Children engage in conversation with their caregivers from their second year of life and reciprocal social interaction even before that (Bruner, 1985). Further, children's language development is predicted most strongly not by the amount of language they hear, but rather by the amount of conversational turns they engage in (Romeo et al., 2018). Given that young children struggle with informativity and perspective-taking, how do conversations with caregivers succeed? One possibility is that parents naturally adopt some of the strategies used by experimenters in training studies of conversational pact formation: Pointing out when references are ambiguous, clarifying when children have produced insufficient information, and/or helping children to conceptualize their referential targets in better ways. Indeed, parents' contingent responses to children's incorrect or incomplete utterances have been implicated in language learning more broadly (Chouinard & Clark, 2003).

Parents spend a tremendous amount of time with children, and often have quite accurate models of their children's lexical and conceptual knowledge (Fenson et al., 1994). Prior work has shown that parents leverage their knowledge of their children's vocabularies to calibrate the informativeness of their referential expressions (Leung, Tunkel, & Yurovsky, 2019). Do they also leverage this knowledge to scaffold the formation of conversational pacts with children? In this study, we asked parents and their 4-, 6-, and 8-year-old children, as well as pairs of adults, to play an adapted version of Krauss and Glucksberg's (1969) director-matcher game. We analyzed their interactions to ask (1) whether parents' contributions support the formation of conversational pacts, (2) what strategies parents use to support pact formation, and (3)

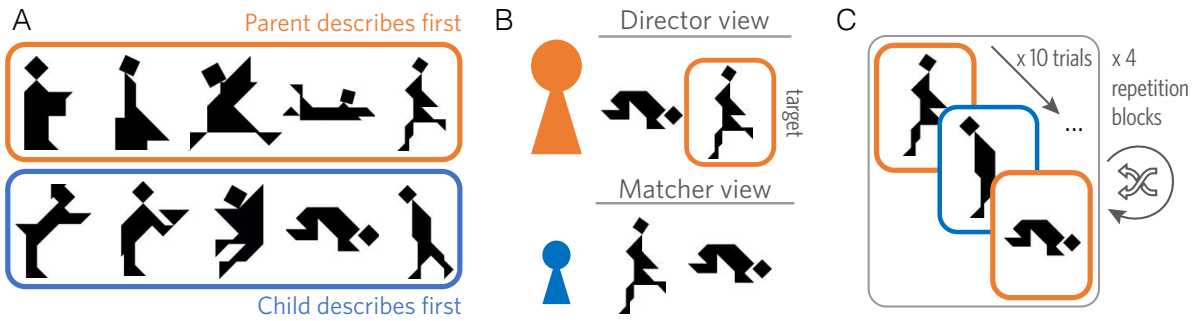

Figure 1: (A) Tangram figures used as referential targets, (B) director and matcher displays, (C) trial sequence.

whether the nature of these pacts changes over development as children acquire more language.

Referential Communication Experiment

Participants Children (ages 4, 6, and 8) and their parents were recruited from a database of families in the local community to achieve a planned sample of 60 parent-child pairs (20 per age group). A total of 75 children and their parents participated. Data from 12 pairs were dropped due to failure to complete the study, leaving a final sample of 63 pairs. For comparison, adult participants were also recruited from a Psychology Department subject pool to achieve a planned control sample of 20 adult-adult pairs. At the time of analysis, only 15 of these pairs had their sessions transcribed.

Stimuli Twelve solid black images of tangrams were normed on Amazon Mechanical Turk for pairwise similarity. Each of the 60 participants made 22 pairwise similarity judgments on a scale from 1-100. Based on these similarity ratings (M = 42.3, SD = 26.1), the ten most dissimilar tangrams were selected for use as stimuli (Fig. 1A).

Design and procedure Pairs of participants were brought into the lab to play a cooperative director-matcher game. They were seated in front of iPads at opposite ends of a table, with a divider preventing them from seeing each other's screens. Participants were told that they would take turns playing *director* and *matcher* roles. The director's task was to describe the *target* image, privately indicated by a blue border, and the matcher's task was to select one of the images on their screen based on the director's description (Fig. 1B). Before beginning the experiment, participants played six practice trials with images of common fruits and vegetables.

The experiment consisted of four repetition blocks of ten trials each (Fig. 1C). Each tangram was the target once per block. We constructed the trial sequence to ensure that participants both alternated roles from trial to trial and alternated roles for each *target* from block to block. For each participant pair, we randomly divided the tangrams into two sets of five: the adult was assigned one set to describe on the first block, and the child was assigned the other set. These sets were interleaved on the first block, such that players alternated roles. On each subsequent block, these sets were swapped such that each tangram was described by each participant exactly twice over the course of the experiment.

On each trial, the target tangram appeared with a foil selected from the set of nine other tangrams. Targets appeared with different foils on different repetition blocks. To ensure that the game would not be too difficult for young children, tangrams most similar to the target (based on similarity norms) did not appear as foils. To discourage participants from using spatial language (e.g. "left side"), the target and foil were shown in randomized order across the two iPads. When the matcher selected an image, it became colorful and a pleasant sound played. Importantly, neither the matcher nor director received feedback about accuracy: the same sound played whether the selection was correct or not.

Pre-processing Sessions were videotaped and subsequently transcribed using Datavyu (2014), an open source coding program. Each video was transcribed and checked by two different coders. Utterances were manually coded as part of a trial or unrelated to the game (e.g. "sit down please"), and unrelated utterances were removed before analysis.

Results

We characterized developmental differences using three measures of communicative behavior. First, we examined accuracy to evaluate whether children were able to succeed at the reference game in collaboration with their parents. Second, we examined conversational turn-taking behavior to evaluate how interactive dialogue may contribute to success. Third, we examined the number of words produced by each partner on each turn to evaluate the efficiency of pacts.

Performance accuracy We began by analyzing task performance across age groups. Because pairs of adults were consistently at ceiling throughout the task, we focused on the performance of parent-child pairs. We constructed a mixed-effects logistic regression predicting whether the matcher successfully chose the correct referent on each trial. The model included (continuous) fixed effects of age and repetition block, random intercepts for each tangram and pair of participants, and random effects of repetition block for each

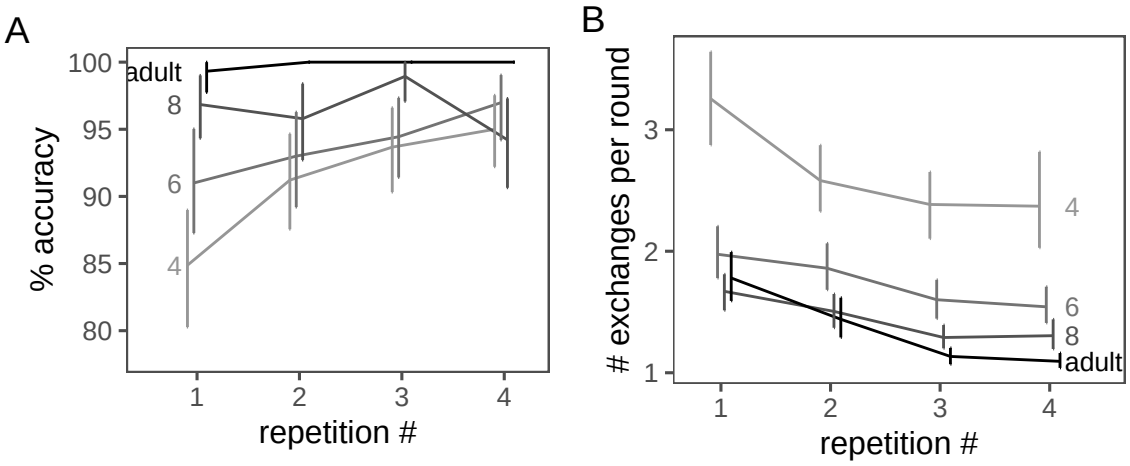

Figure 2: (A) Accuracy improves over time for all groups. (B) Number of dialogue exchanges is higher for younger children. Error bars are 95% CIs. The drop in accuracy for 8-year-olds is likely an artifact of a small set of children losing interest in the task.

pair of participants[1].

Initial accuracy was well above chance for all age groups, indicating that even young children can succeed in this referential task with their parents. We also found a significant main effect of age (β = 0.33, t = 3; $p = .003$): Pairs with younger children performed significantly worse than pairs with older children. Critically, however, accuracy improved significantly over the four repetition blocks for all groups (β = 0.49, t = 3.49, $p < .001$; Fig. 2A). Such improvement for 4-year-olds contrasts with previous results showing no improvement in accuracy with pairs of kindergarteners (age 5, Krauss & Glucksberg, 1969).

Interactive dialogue exchanges If the ability of children of different ages to successfully establish reference depends on scaffolding provided by their parents, we would expect additional dialogue exchanges for younger children. We quantified dialogue by counting the total number of distinct turns of continuous speech on each trial.

We constructed a mixed-effects model predicting the continuous number of exchanges with the same effect structure as reported in the previous section. Consistent with previous work, and replicated in our adult control condition, we found a significant main effect of repetition: Fewer dialogue turns were required on later trials (β = -0.2, t = -4.28; $p <$ .001; Clark & Wilkes-Gibbs, 1986). In line with our predictions, we also found a significant main effect of age (β = -0.33, t = -3.9; $p < .001$). Pairs with 4-year-old children took roughly one additional turn at each point in the experiment than pairs with older children, who more closely resembled pairs of adults (Fig. 2B). This increased level of interactivity between parents and younger children provides an interesting contrast with previous studies showing decreased interactivity between pairs of younger children (Anderson, Clark, & Mullin, 1994).

These lengthier exchanges may reflect efforts by parents to provide and elicit additional clarification or confirmation, or may simply reflect attentional difficulties. As a rough estimate of the content of parents' responses when their child was the director, we counted the number of question marks in the transcript[2]. We found 44.37% of responses to 4-year-olds contained question marks, compared to 27.57% and 31.01% in 6- and 8-year-olds, respectively ($\chi^2(2) = 22.68$, $p < .001$).

Reduction in length of referential expression A key signature of successful communication among adults is an increase in efficiency over repeated reference (Clark & Wilkes-Gibbs, 1986). As pairs form conceptual pacts, they are able to communicate the same meaning using fewer words. Our control sample of adults replicated this classic effect ($\beta = $ -0.24, $t = $ -5.13, $< .001$). Here, we asked whether parents and children of different ages spontaneously reduce their referential expressions in the same way. We define referential expressions as all utterances produced by the director on a given trial, up until a selection is made by the matcher. Because the total number of words produced on a trial is correlated with the number of dialogue exchanges examined above ($r = $ 0.61), we constructed a normalized efficiency measure that controls for additional turns. Specifically, we defined efficiency as the total number of words produced by the *director* divided by the number of dialogue exchanges on that trial.

Because participants in a pair alternated roles, each par-

[1]In principle, we were also interested in whether there were differences in accuracy when the child vs. parent was matcher, but our sample is likely underpowered for such a finer-grained analysis within games.

[2]Because questions were not always annotated with a question mark, this represents a lower bound on the true estimate of questions.

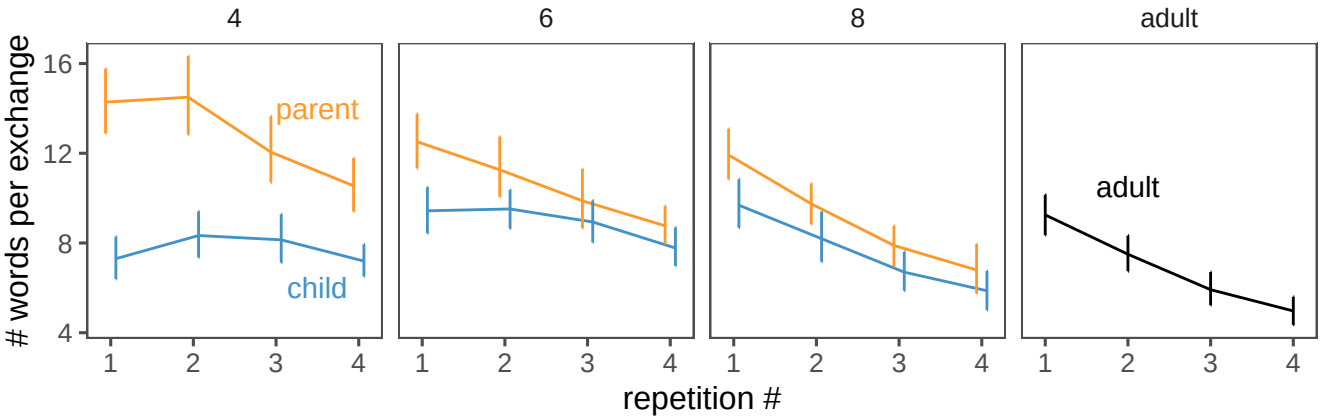

Figure 3: Total number of words in referential expressions produced by children and parents over the course of interaction.

ticipant served as the director twice for each tangram. This structure allowed us to examine how each participant changed their language when they were the director (Fig. 3). Using a mixed-effects model, we predicted the (log) number of words per exchange on each trial, including fixed effects of age, repetition block, and speaker identity (parent vs. child) as well as all of their interactions. We also included random intercepts at the tangram-level and maximal random structure at the dyad level (i.e. intercept, slopes for repetition block and speaker identity, and their interaction; Barr, Levy, Scheepers, & Tily, 2013). All variables were centered to allow interpretation of lower-order terms as effects at the average level of the other terms. We found significant main effects of repetition block (β = -0.13, t = -6.2, $p < 0.001$), speaker identity ($\beta = 0.15, t = 7.09, p < 0.001$), and age ($\beta$ = -0.05, t = -2.06, p = 0.043). All else being equal, directors used fewer words over subsequent repetitions, children used fewer words than their parents, and pairs with older children used fewer words than pairs with younger children. However, these main effects were clarified by several interactions of interest.

First, while parents on average used more words as director than their children did, we found a significant interaction with the child's age (β = -0.04, t = -3.15, p = 0.003). This gap between parent and child utterance length was largest at age 4 but nearly disappeared by age 8. Second, we found that parents reduced their utterance length over time more strongly than children did, holding age group constant (β = -0.03, t = -3.73, $p < 0.001$). Third, having older children in a pair supported stronger reduction overall, (β = -0.04, t = -3.8, $p <$ 0.001).

Finally, an intriguing developmental question is whether reduction changes over the course of development: are 4-year-olds able to become more efficient as common ground is built, in the same way 8-year-olds do? Because parents of different age groups display similar slopes of reduction, this question is addressed by a 3-way interaction. We found that this interaction was not significant ($\beta = 0.01$, $t = 1.8$, p = 0.076), although our sample was likely underpowered to detect this higher-order interaction.

Analyzing the content of pacts

Pairs of age-matched children are notoriously poor at repeated referential communication, failing to coordinate on mutually comprehensible referential expressions (Krauss & Glucksberg, 1977). Our results demonstrate that children and their parents can solve this coordination problem, converging on increasingly accurate and efficient pacts like pairs of adults. What allows children to coordinate with their parents but not with their peers?

One possibility is that children lack the ability to *adapt* to their partner: they have a strong preference for a particular idiosyncratic description and are not sensitive to the possibility that their partner may not understand it. Under this hypothesis, children fail with other children because they each continue to use mutually incomprehensible expressions, and only succeed with their parent as a result of the parent's flexibility. Alternatively, young children may be able to adapt successfully but be unable to generate good initial candidate labels. In this case, pairs of children may fail because neither partner can generate good enough labels to start the pact-formation process, while children and parents succeed because parents seed the first good candidate label.

These accounts make different predictions about who is adapting to who: do pacts originate with children, or with adults? We distinguish these accounts by quantitatively analyzing the natural-language transcripts.

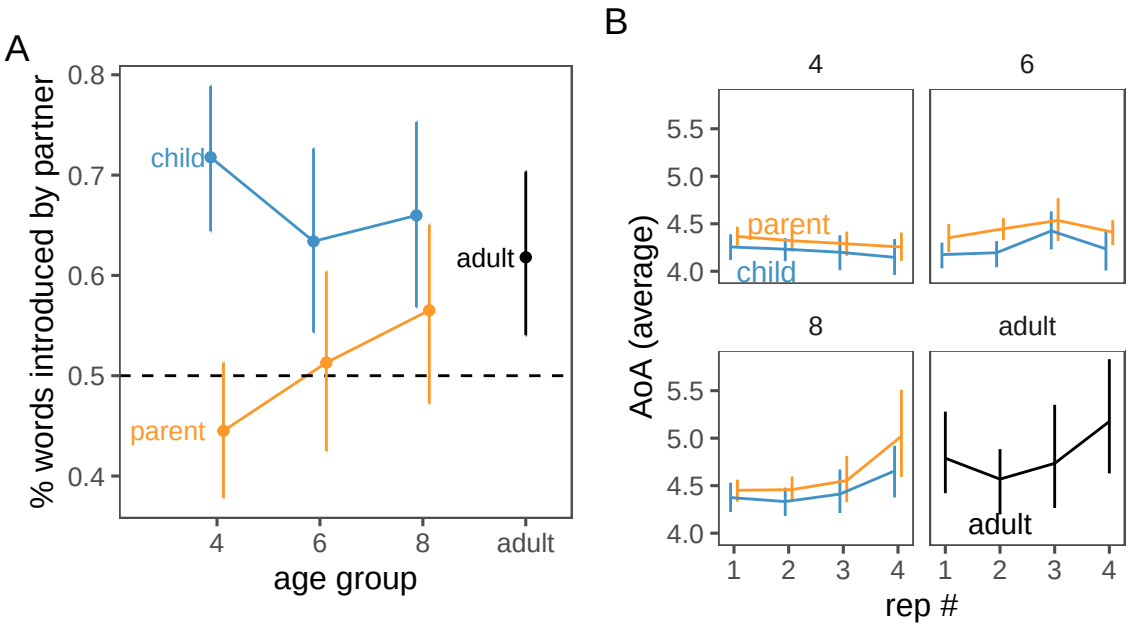

Figure 4: (A) Probability of words used on final round first occuring with child or parent. (B) Complexity of language used by different age groups estimated from words' average age of acquisition. Error bars are 95% CI.

Where do pacts come from?

For each word in the final description of a tangram, we checked whether it had appeared in an earlier referential expression for that tangram. We noted the first trial where it appeared, and who was director when it was produced[3]. The proportion of words originating with the child and parent is shown in Fig. 4A. We observed an asymmetry: the words used by children on the final repetition were more likely to have originated with their parents than the words used by parents were to originate with their children. In addition, this gap appeared to close with older groups, with parents more likely to adopt words introduced by older children.

We tested this hypothesized interaction using a mixed-effects logistic model predicting whether each word appearing on the final repetition for each tangram was introduced by the current director or by their partner. We included fixed effects of age group and the speaker identity (parent or child), as well as random intercepts for each pair of participants and each tangram. We found a significant main effect of speaker identity, with the words used by children more likely to originate with their partner than the words used by parents, (β = 0.37, t = 4.76, $<$ .001). Additionally, we found a weak but significant interaction between speaker and age, indicating that this asymmetry was smaller for older children (β = -0.1, t = -2.19, p = .028). Thus, parents—especially parents of younger children—appear to be the source of the words that persist in successful conceptual pacts.

How do pacts differ across ages?

Given that the labels used by younger children tend to originate with their parents, it is tempting to conclude that adaptation for these pairs is entirely one-sided. However, the resulting pacts may have nonetheless been shaped by children. In particular, we suggest that even though young children may not be able to *supply* good initial labels, they may serve as "simplifiers" or "filters" that constrain the complexity of the resulting pact, since it must pass through their comprehension and production system (Hudson Kam & Newport, 2005).

We tested this hypothesis by asking about the linguistic complexity of the pacts developed in each age group. To estimate the complexity of individual words, we used self-reported age of acquisition (AoA) estimates (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). Our assumption, based on Kuperman et al.'s work, is that more complex words are harder to acquire, and thus have later ages of acquisition (AoA). To estimate the complexity of referential expressions as a whole, we simply averaged the AoAs of all the words they contained[4]. In all age groups, parents produced more linguistically complex referential expressions than children, but as children get older, both their expressions and their parents' become more linguistically complex over repetitions of the same tangram (like adults) (Fig. 4B).

We confirmed this observation with a mixed-effects model

[3]To match different forms of the same word (e.g. "jumping" vs. "jumped") we first lemmatized each word. We also filtered out stop words ("the", "with"), as well as common words that were not part of the pacts ("person", "box"), and excluded words that appeared for the first time on the final repetition of each target.

[4]We first removed any words in a list of 174 stop words like "at" and "me" from the `tm` package (Feinerer, Hornik, & Meyer, 2008)

including all possible interactions of age, person (child vs. parent), and repetition number, and maximal random effect structure as before. We found significant main effects of age ($\beta = 0.06$, $t = 2.51$, $p = 0.015$), and person ($\beta = 0.15$, $t = 4.61$, $p < 0.001$), but not repetition ($\beta = 0.03$, $t = 1.52$, $p = 0.134$). All else being equal, pairs with older children produced more complex utterances, and parents produced more complex utterances than children. Further, we found a significant interaction between age and repetition number, suggesting that later repetitions became more complex for dyads with older children ($\beta = 0.03$, $t = 2.76$, $p = 0.008$). This interaction may be driven by reduction processes, as more distinctive words tend to persist as common words are dropped (Hawkins, Frank, & Goodman, 2020). No other effects were significant, indicating that parents' utterances remained a constant level of complexity above children's even as children's utterances became more complex. These results suggest that pact formation depends not only on parental input, but on children as well—pacts may originate from parents, but children control their complexity.

Discussion

The ability to collaborate with conversational partners on intended meaning is crucial for successful communication. It may be especially important to achieve successful communication with *children*, who regularly encounter novel objects and situations they do not yet have precise words for. In this study, we used a classic repeated reference game to investigate the collaborative processes at play in interactions between parents and their children. While children often struggle to succeed in such games with their peers, we found that even young children were able to successfully form referential pacts with their parents: Accuracy increased, fewer turns were taken, and utterances gradually became shorter over the course of interaction. Furthermore, our results exposed key differences in parental behavior across ages: parents of younger children engaged in more conversational turns, asked more questions, produced longer referential expressions per turn, and adopted less complex labels. These findings indicate that parents may spontaneously adopt some of the strategies used in training studies (e.g. Matthews et al., 2007) to scaffold pact formation with their children.

Our results also contribute to the debate over why children may struggle to form referential pacts with each other. One family of explanations has suggested that children simply cannot take their partner's perspective into account, and produce references only they understand (Krauss & Glucksberg, 1977), or cannot judge which expressions are more informative in the first place, even for themselves (Asher & Oden, 1976; Robinson & Robinson, 1977). An alternative account is that errors may partly arise from children's difficulty in the role of matcher: They may struggle to provide sufficient feedback to their partner or ask for clarification when it is necessary (Anderson et al., 1994).

While we find some support for each of these accounts in our data, our analysis of the asymmetry in the introduction of pacts suggests an additional possibility: Young children may simply be unable to supply a good enough label on their own, possibly due to limited expertise in identifying and labeling the most informative features of novel objects. Once presented with an appropriate label from a parent, however, they will readily adopt and use it. In this way, the processes underlying the mutual "learning" of referential pacts, in the service of coordinating on meaning with a specific partner, may be the same as those underlying language acquisition. Thus the everyday interaction of parents and children trying to understand one another may not only scaffold local pact formation, but support communicative development more broadly.

Stimuli, data, and analysis code available at
`https://osf.io/vkug8/`

Acknowledgements

This research was funded by a James S. McDonnell Foundation Scholar Award to DY.

References

Anderson, A. H., Clark, A., & Mullin, J. (1994). Interactive communication between children: Learning how to make language work in dialogue. *Journal of Child Language*, *21*(2), 439–463.

Asher, S. R., & Oden, S. L. (1976). Children's failure to communicate: An assessment of comparison and egocentrism explanations. *Developmental Psychology*, *12*(2), 132.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(6), 1482–1493.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111–114.

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, *30*(3), 637–669.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Datavyu Team. (2014). Datavyu: A video coding tool. *Databrary Project, New York University*. Retrieved from `http://datavyu.org`

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*(2), 159–184.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in r. *Journal of Statistical Software*, *25*(5), 1–54.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in early

communicative development. *Monographs of the Society for Research in Child Development*, i–185.

Hawkins, R. D., Frank, M. C., & Goodman, N. D. (2020). Characterizing the dynamics of learning in repeated reference games. *Cognitive Science*.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195.

Krauss, R. M., & Glucksberg, S. (1969). The development of communication: Competence as a function of age. *Child Development*, 255–266.

Krauss, R. M., & Glucksberg, S. (1977). Social and nonsocial speech. *Scientific American*, *236*(2), 100–105.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, *44*(4), 978–990.

Leung, A., Tunkel, A., & Yurovsky, D. (2019). Parents calibrate speech to their children's vocabulary knowledge. In *Proceedings of the 41th annual conference of the cognitive science society* (pp. 651–656).

Matthews, D., Lieven, E., & Tomasello, M. (2007). How toddlers and preschoolers learn to uniquely identify referents for others: A training study. *Child Development*, *78*(6), 1744–1759.

Robinson, E. J., & Robinson, W. P. (1977). Children's explanations of communication failure and the inadequacy of the misunderstood message. *Developmental Psychology*, *13*(2), 156–161.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710.