American Psychological Association

Developmental Psychology

0012-1649/19/$12.00

2019, Vol. 55, No. 5, 951–966
http://dx.doi.org/10.1037/dev0000693

Interactive Contexts Increase Informativeness in Children's Referential Communication

Myrto Grigoroglou and Anna Papafragou
University of Delaware

Adults adjust the informativeness of their utterances to the needs of their addressee. For children, however, relevant evidence is mixed. In this article we explore the communicative circumstances under which children offer informative descriptions. In Experiment 1, 4- and 5-year-old children and adults described a target event from a pair of almost identical events to a passive confederate listener who could either see or not see the referents. Adults provided disambiguating information that picked out the target event but children massively failed to do so (even though 5-year-olds were more informative than 4-year-olds). Furthermore, both children and adults were more likely to mention atypical than typical disambiguating event components. Because of the contrastive nature of the task, the listener's visual access had no effects on production. Experiment 2 was a more interactive version of Experiment 1 where participants played a guessing game with a "naïve" listener. In this context, children (and adults) became overall more informative, and the difference between child groups disappeared. We conclude that the informativeness of children's event descriptions is heavily context-dependent and is boosted when children engage in a collaborative interaction with a "true" interlocutor.

Keywords: reference, informativeness, event cognition, instruments, collaboration

Successful communication requires people to take into account the informational needs of their communicative partners. Consider, for example, a situation where two people are driving on the highway and the driver says to the passenger "Look at the bird!" Unless the speaker specifies which bird he is talking about (e.g., "the bird on your right"), the listener would most likely not be able to find the bird the speaker is referring to. Adults readily make such adjustments to their listeners' informational needs in order to facilitate comprehension (e.g., Brennan & Clark, 1996; Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Horton & Keysar, 1996; Lockridge & Brennan, 2002; Nadig & Sedivy, 2002).

For children, however, evidence for speaker adjustments is mixed. Several studies have reported that children, even at a very young age, can successfully adjust the informational content of their referential devices to the knowledge of their listener (e.g., Bahtiyar & Küntay, 2009; Matthews, Lieven, Theakston, & Tomasello, 2006; Nadig & Sedivy, 2002; Nilsen & Graham, 2009; O'Neill, 1996). For example, in studies where children instructed an adult about how to manipulate objects in a visual display, children aged 3- to 6-years-old were more likely to use a scalar adjective (e.g., "Pick up the *little* glass") to disambiguate between two similar objects of different size (e.g., little vs. big glass) when their adult partner could see both of these objects than when the partner could only see one of the objects (Nadig & Sedivy, 2002; see also Bahtiyar & Küntay, 2009; Nilsen & Graham, 2009). Other studies have indicated limitations in children's ability to provide descriptions that match their listener's informational needs (e.g., Davies & Katsos, 2010; Deutsch & Pechmann, 1982; Girbau, 2001; Perner & Leekam, 1986; Sonnenschein & Whitehurst, 1984). For instance, in several referential communication studies 2- to 8-year-old children frequently produced underinformative descriptions of referents when addressing ignorant interlocutors (e.g., Davies & Katsos, 2010; Deutsch & Pechmann, 1982; Girbau, 2001; Sonnenschein & Whitehurst, 1984). Even in studies where children showed sensitivity to a listener's visual perspective when they produced referential descriptions, 5-year-olds appropriately disambiguated between two almost identical referents less than half of the time (e.g., Bahtiyar & Küntay, 2009, Exp.1; Nilsen & Graham, 2009).

At present, the factors that lead to children's success or failure with establishing reference await a full synthesis. The current study aims to reconcile divergent findings in prior work by probing the circumstances under which children provide informative referential descriptions in accordance with a listener's needs. To preface our investigation, we discuss possible sources of children's difficulties with reference production and motivate the current paradigm.

Sources of Children's Difficulties With Reference Production

Reference production is a remarkably complex process that involves the coordination of visual, social, and linguistic information. Successful disambiguation requires speakers to scan the visual context and distinguish the features that set apart the target

This article was published Online First February 11, 2019.

Myrto Grigoroglou, Department of Linguistics and Cognitive Science, University of Delaware; Anna Papafragou, Department of Psychological and Brain Sciences, University of Delaware.

Correspondence concerning this article should be addressed to Myrto Grigoroglou, who is now at Department of Applied Psychology and Human Development, University of Toronto, 252 Bloor Street West, Toronto, ON M5S 1V6, Canada. E-mail: m.grigoroglou@utoronto.ca

from other possible referents in view. Furthermore, speakers have to consider what information is shared or not with their conversational partner to plan utterances that are as informative as required by their partner's knowledge and the purpose of the exchange (see maxim of quantity; Grice, 1975). Finally, speakers need to integrate all this information into well-formed utterances. Based on prior work, one could argue that children have problems at each of these stages of reference production.

A first line of work has proposed that children's problems with reference arise because children do not understand that referring relies on finding differences between a target referent and other objects in view (e.g., Deutsch & Pechmann, 1982; Matthews, Lieven, & Tomasello, 2007, 2012; Nilsen & Mangal, 2012; Whitehurst, 1976; Whitehurst & Sonnenschein, 1981). In support of this proposal, clarification questions by the listener that highlight the contrast between the target and contextual alternatives (e.g., "Which big ball?") help children produce more informative repairs of their original, underinformative descriptions (Bacso & Nilsen, 2017; Deutsch & Pechmann, 1982; Matthews et al., 2012, 2007; see also Golinkoff, 1986; Nilsen & Mangal, 2012; Uzundag & Küntay, 2018). In further support of this position, children tend to produce informative descriptions in tasks where they have to contrast two potential referents that differ only in one dimension (e.g., size; see Nadig & Sedivy, 2002) but provide ambiguous, underinformative descriptions in studies where they are presented with more than two potential referents that differ along more than one dimension (e.g., size, color and shape; see Deutsch & Pechmann, 1982; Ford & Olson, 1975). It should be noted, however, that children frequently produce underinformative referential descriptions even in contexts where there are only two contrasting objects that differ in only one feature and, in principle, the contrast should be highly salient (e.g., Davies & Katsos, 2010; Nilsen & Graham, 2009; Rabagliati & Robertson, 2017). Relatedly, eye-tracking studies with children and adults have shown that, despite the presence of a link between visual search (i.e., gazes at the referential competitor) and referential informativeness, eye-gaze patterns do not always predict informativeness (Brown-Schmidt & Tanenhaus, 2006; Davies & Kreysa, 2017, 2018; cf. Rabagliati & Robertson, 2017). Thus, it is unclear whether children fail to identify a target referent because they do not notice the differences between the target and its competitors or because they do not realize how these differences are relevant for the listener's successful identification of the target.

A second line of research suggests that children's successes and failures with informativeness may be explained by differences in the communicative circumstances embedded into experimental paradigms. For instance, children appear more likely to take into account their addressee's perspective in tasks where the addressee is a "real" interlocutor, either a parent (O'Neill, 1996; O'Neill & Topolovec, 2001) or a person other than the (primary) experimenter (Bahtiyar & Küntay, 2009; Nadig & Sedivy, 2002; Nilsen & Graham, 2009). In contrast, children appear to be less informative when asked to disambiguate referents for a fictional character displayed on a computer screen (Davies & Katsos, 2010), an imaginary addressee in a pretend conversation (Girbau, 2001), or the experimenter herself (Davies & Kreysa, 2018; Rabagliati & Robertson, 2017). Thus, the profile of the listener seems to affect children's informativeness. Although it is not clear exactly what aspect of the listener's profile is responsible for children's successes and failures, there is reason to suspect that, in tasks with "real" addressees, the expectation of the listener's communicative reaction shaped children's production. Adult studies suggest that speakers are more likely to mention extra details when retelling a joke to attentive listeners as opposed to distracted listeners (Kuhlen & Brennan, 2010) or when retelling stories to naïve interlocutors as opposed to the experimenter's confederates (Brown & Dell, 1987; Lockridge & Brennan, 2002; see also Brennan, Galati, & Kuhlen, 2010). In this study, we examine how communicating with an active/collaborative versus a passive listener affects children's referential production. To the best of knowledge, no other study has tested directly how the listener's profile affects children's referential choices.

Finally, it is possible that children's underinformativeness relates to limitations in the process of language production itself (e.g., see Adams & Gathercole, 2000; Levelt, 1989; McDaniel, McKee, & Garrett, 2010). Even though not the focus of the present work, several studies share the assumption that linguistic (Ford & Olson, 1975; Nilsen & Graham, 2009, 2012) and cognitive abilities (e.g., working memory, executive functioning; Bacso & Nilsen, 2017; Nilsen & Graham, 2009; Nilsen, Varghese, Xu, & Fecica, 2015; Wardlow & Heyman, 2016; Uzundag & Küntay, 2018) mobilized during production are, to some extent, linked to referential informativeness and can explain the increase of informative contributions over development (see Matthews, Biney, & Abbot-Smith, 2018, for a recent review).

The Present Study

In this study, we examine children's informativeness for purposes of reference resolution. Unlike past work that has focused on nominal reference, we focus on reference to events. The linguistic and nonlinguistic representation of events is a novel and fast-growing area of linguistic and cognitive studies (e.g., Bunger, Skordos, Trueswell, & Papafragou, 2016; Bunger, Trueswell, & Papafragou, 2012; Hartshorne, Pogue, & Snedeker, 2015; Lakusta & Landau, 2012). Reference production for events is more complex than nominal reference because events may have multiple participants, each of whom might be described at various levels of detail. Here, in two experiments, we asked whether 4- and 5-year-old children (and adults) are able to successfully identify a target event referent from a pair of closely matched events that differ in only one feature. We chose such contrastive contexts to highlight the difference between the target referent and the distractor, because, as mentioned already, prior work suggests that children often fail to provide an adequately informative description of the target referent because they cannot find the difference between the target and the distractor. If so, such simple, contrastive contexts should make the task of finding the relevant differences easier for children.

Our test items involve event reference where disambiguating information concerns the *instrument* used to perform the event. Linguistically, instruments, unlike agents or patients, are not considered arguments of the verb (e.g., Koenig, Mauner, & Bienvenue, 2003; cf. Rissman, Rawlins, & Landau, 2015) and are typically encoded in a nonobligatory adjunct (i.e., Ving *with a Y*). Cognitively, instruments have been shown to be relatively less salient in event representation (Wilson, Papafragou, Bunger, & Trueswell, 2011). Given that instrument information appears to be encoded selectively in both linguistic and nonlinguistic representations, we were interested in whether children and adults would utilize this information for the demands of a referential task. Our control items included cases where disambiguating event information rested on other features, such as the *object* on which the action is performed, typically encoded in an obligatory noun phrase (i.e., Ving *a Y*) or the *location* where the target event takes place, typically encoded in a nonobligatory adjunct (i.e., Ving *next to a Y*).

Our main goal was to test the idea that the communicative context may affect children's informativeness. Given the great variability in communicative circumstances across tasks in prior work, the effect of this factor to children's informativeness remains unclear. In Experiment 1, participants were asked to identify one referent event from a contrastive pair of events for a passive addressee. In Experiment 2, children had to perform the same task but in the context of an interactive guessing game played with a "naïve" addressee who had to find the "right" picture. If children's informativeness depends on their engagement in a collaborative interaction with an actively involved interlocutor with clear informational needs, children should be more likely to mention appropriate disambiguating information in Experiment 2 compared with Experiment 1.

In addition to the communicative circumstances of the task, we explored two further factors that might shape children's referential choices in production. The first factor was the listener's visual perspective. Similarly to prior referential paradigms, we manipulated the listener's visual access to the events to probe whether participants would be more likely to provide successful disambiguation when their listener did not know which of the two events was the target. The second factor was the typicality of event components for test items. Previous work has shown that instrument typicality affects adults' syntactic choices in production: In retelling a story, adults were more likely to mention atypical instruments ("Steve sliced the bread *with a butter knife*") compared with typical instruments ("Steve sliced the bread *with a bread knife*;" Brown & Dell, 1987; Lockridge & Brennan, 2002). Here we asked whether the predictability/typicality of instruments affects children's linguistic choices.

Our studies examine referential abilities in 4- and 5-year-old children (and adults). The reason we chose these age groups was twofold. First, prior literature on nominal reference where the typical sample was between 4 and 6 years has produced conflicting findings (e.g., Bahtiyar & Küntay, 2009; Davies & Katsos, 2010; Nadig & Sedivy, 2002; Nilsen & Graham, 2009; and previous section). We thus wanted to revisit referential abilities in these age groups and extend them to the domain of event reference. Second, we were interested in whether referential abilities change between 4 and 5 years of age as well as between these early years and adulthood, and if so, whether changes are mostly driven by the nature of the referential task, the listener's access to the event or the typicality of event components.

Experiment 1

Method

Participants. Participants were 96 children and 30 adults. The children fell into two age groups: 4-year-olds (n = 48, 23 girls, mean age of 4;7, range: 4;0–5;0) and 5-year-olds (n = 48, 20 girls, mean age of 5;7, range: 5;1–6;0). All children were monolingual speakers of English and were recruited at daycares in Newark, DE, from a mostly middle-income Caucasian population. Adults were undergraduate students at the University of Delaware and received course credit for their participation. Approval for testing these participants had been obtained from the University of Delaware Institutional Review Board (project title: "The Interface Between Language and Spatial Cognition," protocol number: 165481).

Materials. Materials consisted of pairs of events constructed out of clipart pictures and displayed on a computer screen. There were eight pairs of test events (see Appendix A for full list). Within each pair, the events were identical with the exception of the instrument used to perform an action; furthermore, within a pair, one event included a typical and the other an atypical instrument (e.g., a woman sweeping the floor with a broom vs. a tree branch; see Figure 1). Typicality of instruments was independently rated by a separate group of 14 adults and 16 4- to 5-year-old children. These groups were given a questionnaire (administered orally for the children) about the instruments used to perform everyday actions (i.e., "What do we use to sweep the floor?," etc.). On the basis of participants' responses, we then selected eight events to be used in the main experiment. Overall, for these eight events, adults mentioned the selected typical instruments in 73% of their responses and children in 63% of their responses. For atypical versions of the same events, we chose instruments that were either not mentioned at all or mentioned very infrequently (less than 6% of the time) by both children and adults. The experiment also included eight pairs of control events (see Appendix B for full list). Within each pair, events were also identical with the exception of one disambiguating feature. In four pairs, this feature was the object affected by the action (e.g., a baby holding a rattle vs. the same baby holding a kite; Figure 1). In the remaining four control pairs, the disambiguating feature was the location of the action (e.g., a man watering flowers next to a tree vs. the same man watering flowers but without a tree present; Figure 1).

For purposes of the referential task, one member of each test or control event pair was placed within a red circle that marked the target event (see Figure 1). For test events, we created two basic presentation lists that differed only in terms of which member of each pair of the test events was circled. Within each list, half of the time the circle was placed around a typical and the other half around an atypical instrument version of test events. Furthermore, the position (left–right) of the typical and atypical instrument versions were counterbalanced within each list. The control events were then added to the two lists. For control events, the circle was always placed around a specific member of each event pair (for disambiguating location trials, this was the event containing the location object) so control events did not differ across lists. Finally, the two basic presentation lists (each with 16 event pairs) were reversed for a total of four lists.

Procedure. The experimenter introduced participants to her "friend" (a confederate listener) and pointed out to them the display of events on a computer screen. The experimenter gave the participants the following instructions: "These are two twins. They are each doing something different. Look at both twins and tell [the "friend"] what the twin inside the circle is doing. She has a picture of the twins too, but she doesn't know which one we are talking about." Participants saw that the listener had a binder which contained color printouts of the pictures on their computer screen but lacked circles around the target pictures.

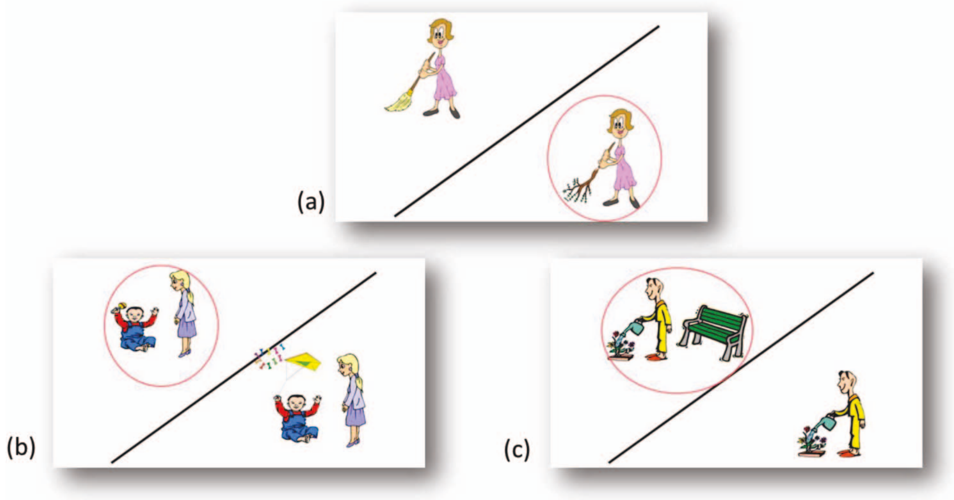

Figure 1. Sample stimuli for Experiment 1. An example of a test stimulus is (a), where disambiguating information for the target appears in the instrument (a girl is sweeping with a tree branch). Examples of control stimuli are (b) and (c), where disambiguating information appears in an object (a baby is holding a rattle) or location (a man is watering flowers next to a bench), respectively. See the online article for the color version of this figure.

Participants were randomly assigned to one of two conditions. In the visual access condition, participants and listener were seated next to each other, so that they both had visual contact with the pictures described. The listener looked at the screen as participants were describing the events and followed along by turning the pages in her binder. In the no visual access condition, the listener was seated across from the participants, so that she could not see the computer screen. In that condition, the listener avoided eye contact with the participants but kept looking into her binder and followed the descriptions, turning the pages as appropriate.

At the beginning of the session, participants performed one practice trial that also consisted of a contrastive pair of events (i.e., a man drawing a dragon vs. the same man drawing a circle). If the participant gave an underinformative description (e.g., "He is drawing"), the experimenter provided appropriate feedback (e.g., "See? It's the man drawing the dragon!"). This was the only corrective feedback children received during the study.

Coding

Participants' descriptions were coded for the presence of disambiguating information. For test items, we coded whether participants mentioned the disambiguating instrument. Mention types were classified as *Instrument Prepositional Phrase/Verb* when the instrument appeared in a phrase introduced by the preposition *with* or the verb *using* (e.g., "The woman sweeping with a broom/using a broom") or was incorporated into the verb (e.g., "She is brooming") and *Other mention* (e.g., "She is holding a broom;" "There is a broom"). For control items, we noted whether speakers mentioned the disambiguating object or location. For objects, mention types were either *Object Noun Phrase* (e.g., "The baby is shaking a rattle") or *Other mention* (e.g., "The baby is playing with a rattle"). For locations, mention types were either *Location Prepositional Phrase* (e.g., "He is watering flowers next to a bench") or *Other mention* (e.g., "There is a bench").[1]

Results

Data analytic strategy. We measured the mention of disambiguating information for test items (typical and atypical instruments) and control items (objects and locations). This measure was a binary outcome variable (target present = 1, target absent = 0). The analysis dataset for Experiment 1 consisted of 126 subjects × 16 items (eight test, eight control) = 2,016 observations. Inspection of the data showed five missing cases (0.2%, all control items). Data were analyzed using multilevel logistic mixed-effects modeling with crossed random intercepts for participants and items (Baayen, 2008; Baayen, Davidson, & Bates, 2008). This analytical approach is ideal for our study because it allows for participants and items to be treated as random factors in a single model and, at the same time, it is the indicative treatment of categorical data (cf. Barr, 2008; Jaeger, 2008). All models presented in this section were fit using the *glmer* function of the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in the R Project for Statistical Computing (R Core Team, 2015). The mention of target disambiguating information was analyzed in separate models for test and control items.

Mention of target disambiguating information. Beginning with test items, we analyzed the data using a model that included Mention of Instruments as the binary dependent variable and participants and items as crossed-level random intercepts. Figure 2 summarizes the data. The best fit for these data was a model that included typicality (typical, atypical) as a first-level fixed predic-

[1] For both test and control items, cases where children simply pointed at the target without verbal disambiguation were coded as *No mention*.

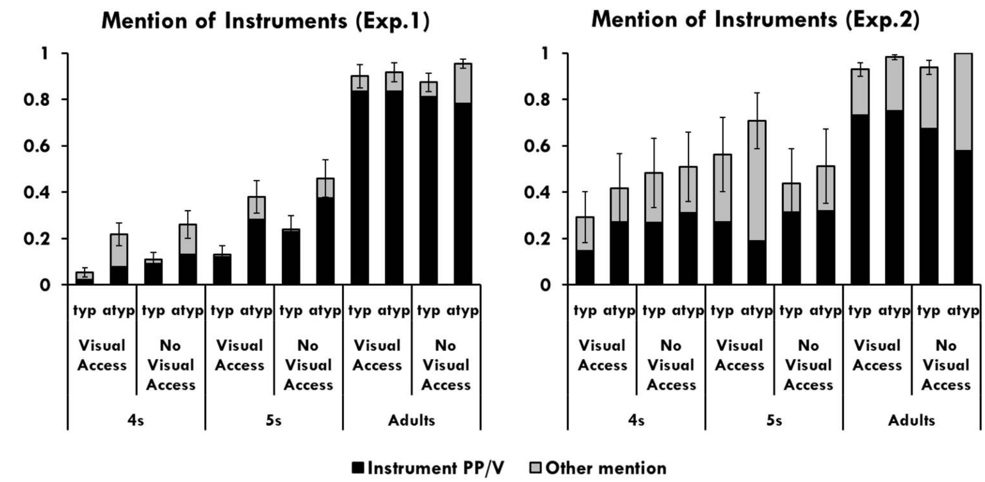

Figure 2. Proportion mention of typical and atypical instruments per age group and visual access in Experiments 1–2 (split by different types of mention). Error bars represent standard error for all mentions combined.

tor, age (4-year-olds, 5-year-olds, adults) as a second-level predictor, and their interaction. Visual access (visual access, no visual access) did not significantly improve model fit based on a chi-square test of the change in −2 restricted log likelihood and was, therefore, not included in the final model, $\chi^2(1) = 2.38$, $p = .12$.[2] The fixed effect of typicality was coded with centered contrasts (−.5, .5) and the fixed effect of age was analyzed with two simple contrasts comparing adults with children (c_1: −.66, .33, .33) and 5-year-olds to 4-year-olds (c_2: 0, −.5, .5). The same coding strategy was followed in all the following analyses. Table 1 presents the parameter estimates for the multilevel model of instrument mention. The model revealed a significant effect of typicality: Atypical instruments were mentioned more often than typical instruments ($M_A = .47$, $M_T = .32$). There was also a significant effect of age: Adults mentioned instruments more frequently than children ($M_A = .93$, $M_C = .23$), and 5-year-olds more frequently than 4-year-olds ($M_5 = .30$, $M_4 = .15$).

Turning to control items, Figure 3 summarizes the data. We analyzed these data using a model that included *mention of disambiguating feature* as the binary dependent variable and participants and items as crossed-level random intercepts. The best fit for the data was a model that included disambiguating feature (objects, locations) as a first-level predictor, age (4-year-olds, 5-year-olds, adults) as a second-level predictor, and the interaction between them. Visual access did not significantly improve model fit (based on a chi-square test of the change in −2 restricted log likelihood) and was, thus, not included in the final model, $\chi^2(1) = .19$, $p = .67$. Similarly to the previous analysis, the fixed effect of disambiguating feature was coded with centered contrasts (−.5, .5) and the fixed effect of Age with simple contrasts. Table 2 presents the parameter estimates for the multilevel model of disambiguating feature mention. The model revealed a significant effect of age and a significant interaction between age and disambiguating feature. Adults were more likely to mention disambiguating features than children ($M_A = .97$, $M_C = .31$) and 5-year-olds more likely than 4-year-olds ($M_5 = .38$, $M_4 = .23$). The interaction was due to the fact that children mentioned disambiguating objects more frequently than disambiguating locations ($M_O = .36$, $M_L = .24$) but adults mentioned both types of disambiguating features equally frequently ($M_O = .95$, $M_L = .99$).

Discussion

Experiment 1 showed that 4- and 5-year-old children, unlike adults, failed to produce event information that would help a communicative partner disambiguate the target referent. This result suggests that visual contrast by itself is not sufficient to successfully clarify what information is relevant and should be used to disambiguate a referent in a given context (cf. also Davies & Katsos, 2010; Nilsen & Graham, 2009). Specifically, adults consistently mentioned disambiguating instruments in their event descriptions but 4- and 5-year-old children in our sample did so very infrequently. Similarly, adults mentioned disambiguating objects and locations equally frequently and almost all the time. Children, however, were less likely to include these features in their descriptions, and when they did, they mentioned objects more frequently than locations. Overall, as children got older, they tended to provide more informative descriptions of events. These results are consistent with and extend a large body of developmental work on nominal reference showing that preschool children frequently produce underinformative utterances (e.g., Deutsch & Pechmann, 1982; Nilsen & Graham, 2009; Sonnenschein & Whitehurst, 1984).

Two further aspects of our findings are worth pointing out. First, both preschoolers and adults in our experiments were more likely to mention atypical/unusual as opposed to typical/common instruments to disambiguate target events. This result generalizes prior findings that involved mostly adults (see Brown & Dell, 1987; Lockridge & Brennan, 2002; cf. also Papafragou, Massey, &

[2] We also performed a Bayesian assessment of the null hypothesis (i.e., that there is no effect of visual access), using Bayes factors in the *brms* package in R (Bürkner, 2017). This analysis showed that the estimated Bayes factor in favor of H_1 (i.e., a model including visual access as a factor) over H_0 (i.e., a model without visual access) was 0.84, suggesting weak evidence for H_0 (Jeffreys, 1961; Raftery, 1995).

Table 1
Parameter Estimates for Instrument Mention in Experiment 1

Effects	Estimate	*SE*	*z*
Intercept	−.21	.28	−.73
Typicality (Typical vs. Atypical)	1.29	.24	5.35^{***}
Age (Adults vs. Children)	−5.24	.48	-10.76^{***}
Age (5- vs. 4-year-olds)	−1.31	.39	-3.39^{***}
Typicality (Typical vs. Atypical): Age (Adults vs. Children)	1.12	.59	1.90
Typicality (Typical vs. Atypical): Age (5- vs. 4-year-olds)	.03	.47	−.07

*** $p < .001$.

Gleitman, 2006). Second, the addressee's visual access to the events did not affect referential choices in any age group. One possible explanation is that participants were completely insensitive to the perspective of the addressee and adopted a fully egocentric model of referential communication. Alternatively, participants may have computed the perspective of the addressee but the communicative pressures of the task (help the addressee find the right picture, independently of whether the addressee has visual access or not) may have prevailed. A third, perhaps more likely possibility, given that adults were overall highly informative and children overall highly underinformative, is that the first of these explanations applies to children and the second to adults. We return to these possibilities in the next experiment.

Experiment 2

Experiment 2 made two main modifications to Experiment 1. First, the addressee was no longer a confederate acting as a passive listener but an actively involved interlocutor introduced as a "naïve" partner. Second, and relatedly, the communicative task had a specific goal (success in a guessing game) that was shared between the participant and addressee and required their coordinated effort.

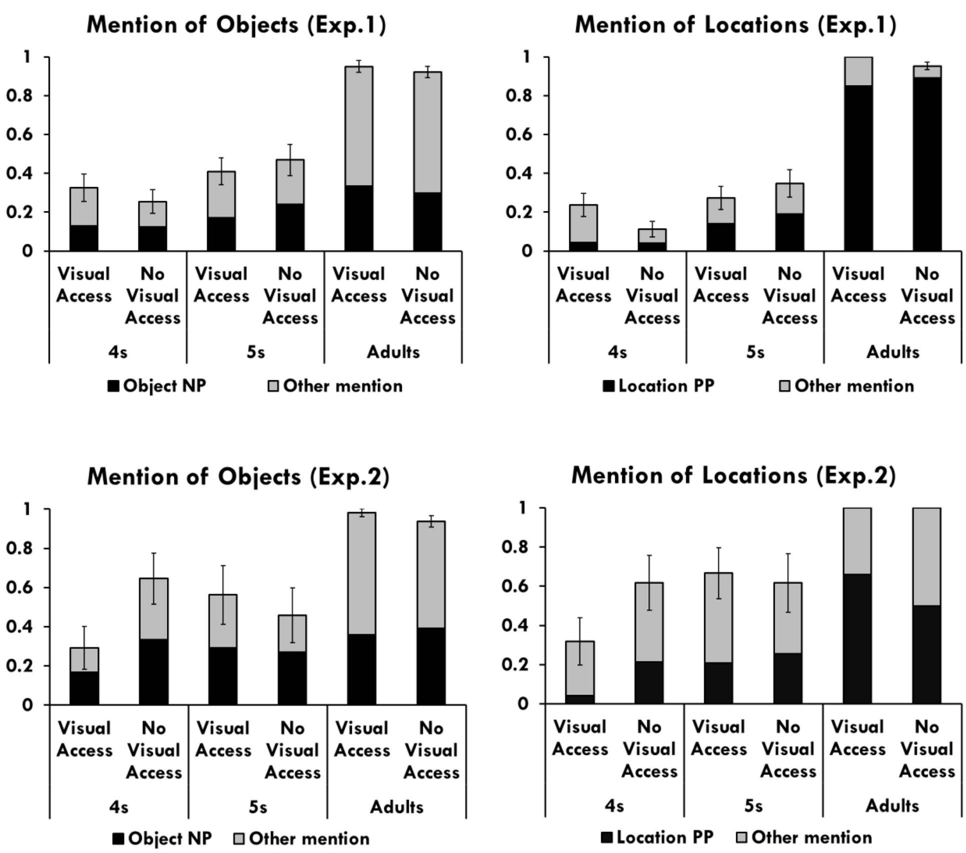

Figure 3. Proportion mention of disambiguating objects and locations per age group and visual access in Experiments 1–2 (split by different types of mention). Error bars represent standard error for all mentions combined.

Table 2
Parameter Estimates for Disambiguating Feature Mention in Experiment 1

Effects	Estimate	*SE*	*z*
Intercept	.77	.39	1.99*
Feature (Object vs. Location)	−.11	.72	−.16
Age (Adults vs. Children)	−6.26	.71	−8.84***
Age (5- vs. 4-year-olds)	−1.04	.35	2.95**
Feature (Object vs. Location): Age (Adults vs. Children)	3.28	1.15	2.85**
Feature (Object vs. Location): Age (5- vs. 4-year-olds)	.10	.40	.25

* $p < .05$. ** $p < .01$. *** $p < .001$.

We reasoned that participants might become more informative in a highly interactive paradigm in which they had to collaborate with an active addressee that had salient informational needs. Although we used a confederate listener, our design incorporated several factors that make confederate use naturalistic (see Kuhlen & Brennan, 2013). To ensure that the listener's needs were as genuine as possible we allowed the listener to interact with the participant freely during most of the introductory phase, kept the scripted utterances to a minimum and incorporated them in a collaborative context where the listener had to take conversational initiative (e.g., ask questions, make a "guess," move the conversation forward).

Method

Participants. Participants were 48 children and 30 adults. None of them had participated in Experiment 1. The children fell into two age groups: 4-year-olds ($n = 24$, 11 girls, mean age of 4;6, range 4;0–5;0) and 5-year-olds ($n = 24$, 10 girls, mean age of 5;5, range 5;1–6;0). All children were recruited at the same local daycares as children in Experiment 1 and were monolingual speakers of English. Adults were undergraduate students at the University of Delaware and received course credit for their participation.

Materials. Materials were identical to Experiment 1.

Procedure. There were two phases in the experiment. In the introductory phase, the experimenter brought the child from her classroom to the testing room. The addressee was instructed to wait outside the testing room and act as if she had just arrived. Experimenter, participant, and addressee were introduced to each other. Unlike Experiment 1, the experimenter and addressee acted as if they did not know each other and the addressee was participating in the study for the first time. The experimenter showed the participant and the addressee where to sit. The participant was seated in front of the computer. The addressee sat either next to the participant, so that they both had visual contact with the pictures described (visual access condition), or right across from the participant, so that the addressee could not see the participant's screen (no visual access condition). Unlike Experiment 1, when testing children, the experimenter showed the participants the pictured pairs of events on the computer screen and asked them if the addressee could see the pictures. Independently of their response, children were instructed to sit on the addressee's chair to make sure that the addressee could or could not see the events (depending on the condition). This ensured that children were aware of the addressee's visual access before the beginning of the test phase. For adults, this step was omitted.

In the main phase, the participant's attention was directed to the pictures on the computer screen and to the circle around the target event. The experimenter showed the participant and the addressee the same binder used in Experiment 1; the binder contained color printouts of the pictures on their computer screen but lacked circles around the target pictures. The instructions to the participant were similar to the instructions of Experiment 1 except for the fact that this was a guessing game where the participant would have to help the addressee: "These are two twins and they are each doing something different, see? [participant's name], your job in this game is to help [addressee's name] find the right picture. To do that, you need to describe what the person inside the circle is doing. Remember that [addressee's name] does not have a circle." In a departure from Experiment 1, the experimenter gave both the binder and some stickers to the addressee and further explained the rules of the game: "[Addressee's name], your job is to guess which picture [participant's name] is talking about. When you make a guess, I want you to put a sticker next to the picture you think is the right one. Ok?" The addressee was explicitly instructed to lift one side of the binder to prevent the participant from seeing what she was doing (in both visual access conditions). Thus, the participants could not see where the addressee placed the sticker. To ensure some communicative authenticity in the addressee's responses, the addressee in this experiment was instructed to react naturally to the experimenter's instructions.

Next, the experimenter told the participant and the addressee that they would practice before the "real game." Practice trials involved more addressee-initiated discourse compared with Experiment 1. At the beginning of the practice trial, the addressee said "I can see two pictures. Which one is it? Tell me about it!" and waited for the participant's response. At the end of the practice trial (independently of the participant's response), the addressee said "I hope I got it right!" and placed a sticker next to the picture that matched the description. If the participant gave an underinformative description, the addressee placed the sticker next to the nontarget picture. The experimenter asked the addressee to show her binder and gave feedback.

The test trials were exactly the same as the practice trial but without any feedback: The participants could not see where the addressee put the sticker until the end of the game. A final difference from Experiment 1 was that, after each test trial, the addressee prompted the participant to move on to the next trial by saying "I am ready for the next one."

Table 3
Parameter Estimates for Instrument Mention in Experiment 2

Effects	Estimate	*SE*	*z*
Intercept	1.66	.39	4.30***
Typicality (Typical vs. Atypical)	.98	.28	3.47***
Age (Adults vs. Children)	−5.47	.88	−6.20***
Age (5- vs. 4-year-olds)	−.87	.78	−1.13

*** $p < .001$.

Results

The analysis dataset for Experiment 2 consisted of 78 subjects × 16 items = 1,245 observations. Inspection of the data showed three missing cases (0.2%, all control items). The same data analytic strategy as in Experiment 1 was used. Beginning with test items, our analysis used a model that included mention of instruments as the binary dependent variable, typicality (typical, atypical) and age (4-year-olds, 5-year-olds, adults), as fixed predictors and random intercepts for participants and items. Visual access and the cross-level interaction between typicality and age did not significantly improve model fit (based on chi-square tests of the change in −2 restricted log likelihood) and, therefore, were not included in the final model, $\chi^2(1) = .03$, $p = .85$ and $\chi^2(1) = 2.59$, $p = .11$, respectively. Table 3 presents the parameter estimates for the multilevel model of instrument mention for Experiment 2. Figure 2 summarizes the data. The model revealed a significant effect of typicality: Atypical instruments were mentioned more frequently than typical instruments ($M_A = .72$, $M_T = .63$). The model also showed a significant effect of age: Adults mentioned instruments more frequently than children ($M_A = .96$, $M_C = .49$) but children groups did not differ ($M_5 = .55$, $M_4 = .43$).

Turning to control items, we analyzed the mention of disambiguating features. Because mention of objects and locations for the adult group was extremely high ($M_O = .96$, $M_L = 1.00$), adult data did not have enough variability and were not included in this analysis. (An assessment of model fit based on chi-square tests of the change in −2 restricted log likelihood for the adult data separately showed that no model other than the empty model with random intercepts for participants and items was a good fit for these data.) We used a model that included fixed effects of disambiguating feature (objects, locations), age (4-year-olds, 5-year-olds), visual access (visual access, no visual access), an interaction between disambiguating feature and age, and an interaction between age and visual access. The model also included crossed random intercepts for participants and items. Table 4 presents the parameter estimates for the multilevel model of disambiguating feature mention for Experiment 2. Figure 3 summarizes the data. The model revealed a significant interaction between disambiguating feature and age: Five-year-olds were more likely to mention locations than objects ($M_O = .49$, $M_L = .63$) but 4-year-olds mentioned objects and locations equally frequently ($M_O = .47$, $M_L = .47$). The model also showed a significant interaction between age and visual access: Four-year-olds were more likely to mention disambiguating features when the listener could not see the events ($M_V = .23$, $M_{NV} = .66$) but 5-year-olds were unaffected by the listener's visual access ($M_V = .64$, $M_{NV} = .49$). There were no other effects or interactions.

Comparison of Experiments 1 and 2

To test for the effect of the interactive paradigm on participants' informativeness we performed a number of comparisons between Experiments 1 and 2.

Mention of Target Disambiguating Information

To compare mention of instruments across the two experiments we used a model that included typicality, age, and experiment as fixed predictors and participants and items as random intercepts (visual access and the three-way interaction between typicality, age, and experiment did not improve model fit and were not included in the final model). Table 5 presents the parameter estimates for the multilevel model of comparing Instrument mention across experiments. Unsurprisingly given the previous analyses, the model yielded significant effects of typicality and age. Importantly, the analysis also showed a main effect of experiment, with participants in Experiment 2 being more likely to mention instruments than participants in Experiment 1 ($M_1 = .39$, $M_2 = .67$).

To compare mention of disambiguating features across the two experiments we used a model that included fixed effects of disambiguating feature, age, and experiment and their two-way interactions (visual access and the three-way interaction between disambiguating feature, age, and experiment did not improve model fit and were not included in the final model). The model also included crossed random intercepts for participants and items. Table 6 presents the parameter estimates for the multilevel model of mention of disambiguating features across experiments. Unsurprisingly, the model yielded a significant effect of age. The model also showed an effect of experiment, with participants of Experiment 2 being more likely to mention disambiguating features than participants of Experiment 1 ($M_1 = .46$, $M_2 = .69$). The model

Table 4
Parameter Estimates for Disambiguating Feature Mention in Experiment 2

Effects	Estimate	*SE*	*z*
Intercept	−.06	.42	−.12
Feature (Location vs. Object)	−.59	.49	−1.22
Age (5-year-olds vs. 4-year-olds)	−.72	.75	−.97
Visual Access (NV vs. V)	.81	.75	1.09
Feature (Location vs. Object): Age (5- vs. 4-year-olds)	1.28	.60	2.13*
Age (5- vs. 4-year-olds): Visual Access (NV vs. V)	3.37	1.52	2.22*

* $p < .05$.

Table 5
Parameter Estimates for Instrument Mention in Experiments 1–2

Effects	Estimate	*SE*	*z*
Intercept	−.18	.28	−.63
Typicality (Typical vs. Atypical)	1.29	.17	7.59***
Age (Adults vs. Children)	−5.29	.44	−12.09***
Age (5- vs. 4-year-olds)	−1.16	.35	−3.30***
Experiment (1 vs. 2)	1.73	.33	5.23***

*** $p < .001$.

further returned an interaction between disambiguating feature and age: Children were more likely to mention disambiguating objects than disambiguating locations ($M_O = .40$, $M_L = .34$) but adults had the reverse tendency ($M_O = .95$, $M_L = 1.00$). Finally, there was also an interaction between disambiguating feature and experiment: Although speakers mentioned both types of disambiguating features more frequently in Experiment 2 compared with Experiment 1 ($M_1 = .46$, $M_2 = .69$), the increase in the mention of disambiguating locations was greater than the increase in the mention of disambiguating objects ($M_{DIFF\_L} = .30$, $M_{DIFF\_O} = .16$).

Additional Measures of Informativeness

To gain a more complete picture of how Experiment 2 affected participants' overall informativeness, we performed two additional analyses. First, we coded participants' descriptions for mention of the agent across all test and control items (cf. Matthews et al., 2006). Even though agents were not relevant for disambiguating the target event, we hypothesized that they would be affected by the nature of the conversational exchange and hence mentioned more in Experiment 2. There were two categories of agent mention: *Agent NP/Pronoun* realized in subject position (e.g., "The lady is brushing her teeth with a toothbrush," "She is brushing her teeth with a toothbrush), and *Other mention* where the person was mentioned but not as agent/subject (e.g., "There is a refrigerator next to *the man*"), or was clearly contrastive ("*That one* doesn't have a rattle"). Figure 4 presents the data.

We used a model that included fixed predictors of age, visual access, and experiment and interactions between age and visual access, and age and experiment (all other interactions did not significantly improve model fit and were, therefore, not included in the final model; all *p*s > 0.05). The model also included random intercepts for participants and items. Table 7 presents the parameter estimates for the multilevel model of comparing mention of agents across experiments. The model yielded a significant effect of age: Overall, adults mentioned agents more frequently than children ($M_A = .94$, $M_C = .56$), but 4- and 5-year-olds did not differ ($M_5 = .61$, $M_4 = .50$). There was also a significant effect of experiment, with participants of Experiment 2 being more likely to mention agents than participants of Experiment 1 ($M_1 = .56$, $M_2 = .65$). These effects were qualified by an Age × Experiment interaction: Children were much more likely to mention agents in Experiment 2 compared with Experiment 1 ($M_1 = .44$, $M_2 = .79$) but no such difference emerged in adults ($M_1 = .93$, $M_2 = .96$). There was also an Age × Visual Access interaction in children: Five-year-olds were more likely to mention agents when the listener did not have visual access to the events ($M_V = .50$, $M_{NV} = .72$) but no such difference emerged in 4-year-olds ($M_V = .56$, $M_{NV} = .45$). No such interaction was found when children as a group were compared with adults (for adults, $M_V = .96$, $M_{NV} = .93$).

Second, we measured the length of participants' responses for all test and control events (i.e., the number of words used for each event description) across experiments. Figure 5 presents the data. This continuous variable was analyzed with a mixed-effects linear model with crossed random intercepts for participants and items. The model was fitted using restricted maximum likelihood of parameters (REML) and included fixed predictors for age and experiment and their interaction (visual access did not significantly improve model fit, $\chi^2 = 2.80$, $p = .09$). Table 8 illustrates the parameter estimates for the multilevel model of comparing length of utterance across experiments. The analysis yielded a significant effect of age and a significant effect of experiment. These main effects were qualified by an interaction between age and experiment: For adults, length of utterance did not differ between the two experiments ($M_1 = 9.03$, $M_2 = 8.57$), but for children, length of utterance was significantly higher in Experiment 2 ($M_1 = 4.26$, $M_2 = 7.61$), with 4-year-olds showing greater gains ($M_1 = 3.78$, $M_2 = 8.23$) than 5-year-olds ($M_1 = 4.74$, $M_2 = 6.98$).

Discussion

Experiment 2 introduced a collaborative exchange with a highly active addressee to the paradigm of Experiment 1. Even though the adult-child gap in informativeness persisted, the difference between 5- and 4-year-olds disappeared, and participants in all age groups were more informative for both test and control items in Experiment 2 compared with Experiment 1. Furthermore, children—but not adults—were more likely to mention agents and to provide

Table 6
Parameter Estimates for Disambiguating Feature Mention in Experiments 1–2

Effects	Estimate	*SE*	*z*
Intercept	1.49	.36	4.20***
Feature (Location vs. Object)	.92	.64	1.43
Age (Adults vs. Children)	−6.59	.68	−9.63***
Age (5- vs. 4-year-olds)	−.95	.34	2.80**
Experiment (1 vs. 2)	1.34	.33	4.04***
Feature (Location vs. Object): Age (Adults vs. Children)	−3.49	1.11	−3.16**
Feature (Location vs. Object): Age (5- vs. 4-year-olds)	−.46	.33	−1.38
Feature (Location vs. Object): Experiment (1 vs. 2)	1.51	.35	4.37***

** $p < .01$. *** $p < .001$.

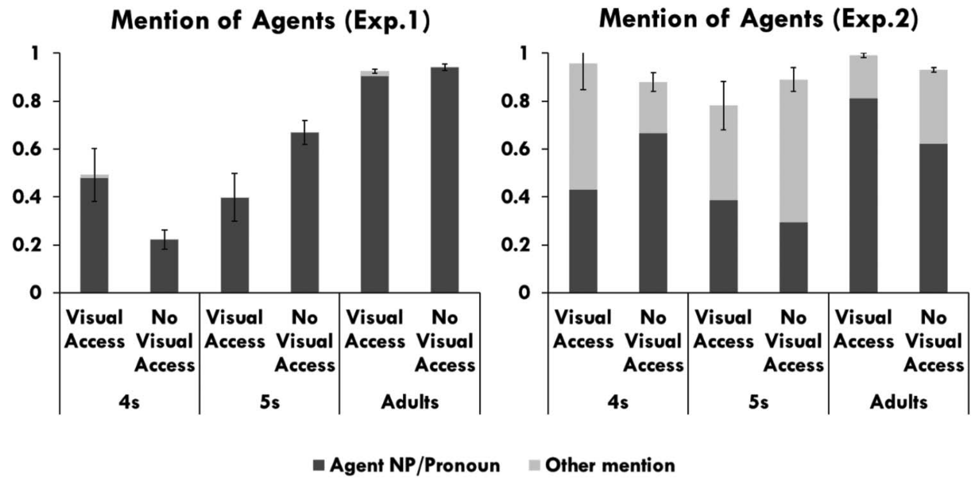

Figure 4. Proportion mention of agents per age group and visual access for Experiments 1–2 (split by different types of mention). Error bars represent standard error for all mentions combined.

longer descriptions in Experiment 2 compared with Experiment 1 (with the increase in utterance length being more pronounced in 4-year-olds). Thus, beyond the narrow focus of providing disambiguating information to pick out the target event, the interactive context of Experiment 2 affected participants' tendency to offer fuller, more informative event descriptions.

Several additional aspects of our data are worth discussing. First, as in Experiment 1, participants in Experiment 2 mentioned atypical instruments more frequently than typical instruments. Second, in Experiment 2, the listener's visual access did not play a role in instrument mention in any age group. However, closer inspection of other parts of the data revealed some evidence for children's sensitivity to visual access: Four-year-olds in Experiment 2 mentioned disambiguating features more often when the addressee could not see the events (even though older children and adults did not show this pattern). Similarly, across experiments, 5-year-olds were more likely to mention agents in the no visual access compared with the visual access condition (adults and younger children showed no such bias). Together with the results from Experiment 1, these findings suggest that children were not completely insensitive to the perspective of the addressee, even though they did not consistently fulfill the addressee's need for information. Even though the current findings do not directly speak to perspective-taking in adults, it is highly likely that adults followed the rules of the game and provided disambiguating information consistently, whether the addressee had access to the target events or not.

Finally, our results reveal differences between types of disambiguating information: Four- and 5-year-old children, across experiments, mentioned objects more frequently than locations but in adults this pattern was reversed. The adult finding might be easier to explain: Disambiguating locations in our stimuli were perceptually more salient than objects, because they were included only in the target event (but were absent in the distractor) and were typically larger in size than the disambiguating objects (see Figure 1). For 4- and 5-year-olds, however, the source of the pattern is less clear. One possibility is that the source is linguistic. Perhaps it was easier for our child participants to plan utterances with (linguistically obligatory) direct object noun phrases than (optional) prepositional phrases. An alternative possibility is that the source of this pattern is conceptual/semantic. Preschoolers in our sample might have preferred to mention event components that were essential to the core event (objects affected by the main action) than conceptually more peripheral features (locations where an action took place). At present, we cannot adjudicate between the two possibilities, although a purely linguistic explanation seems

Table 7
Parameter Estimates for Agent Mention Across Experiments 1–2

Effects	Estimate	*SE*	*z*
Intercept	3.47	.40	8.72***
Age (Adults vs. Children)	−5.24	.73	−7.19***
Age (5- vs. 4-year-olds)	−.84	.80	−1.06
Access (V vs. NV)	.58	.62	.94
Experiment (1 vs. 2)	2.91	.68	4.30***
Age (Adults vs. Children): Access (V vs. NV)	.54	1.37	.40
Age (5- vs. 4-year-olds): Access (V vs. NV)	−4.77	1.52	−3.13**
Age (Adults vs. Children): Experiment (1 vs. 2)	3.98	1.44	2.76**
Age (5- vs. 4-year-olds): Experiment (1 vs. 2)	2.26	1.61	1.41

** $p < .01$. *** $p < .001$.

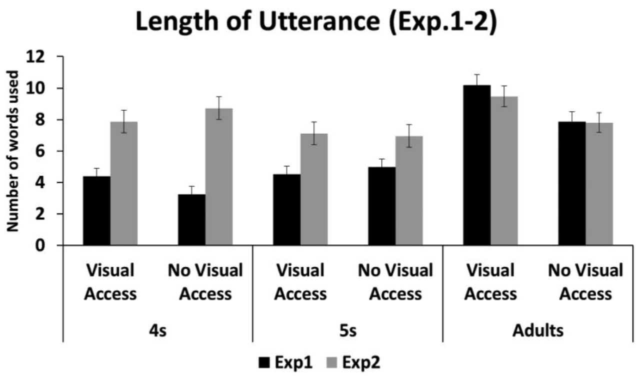

Figure 5. Mean length of utterance per age group and visual access for Experiments 1–2. Error bars represent standard error.

less likely, given that 4- and 5-year-olds in our experiments were too old to have difficulties formulating sentences with PPs and, furthermore, children's (and adults') mentions did not strictly follow the expectation that disambiguating objects would be encoded as direct object NPs and disambiguating locations as adjunct PPs (see Coding section and Figure 3). Regardless of the precise explanation of these findings, it appears that not all kinds of disambiguating information are equally encoded.

General Discussion

The present set of experiments investigated referential production patterns in 4- and 5-year old children and adults, focusing on event reference. Specifically, we explored how various factors such as the communicative circumstances of the task, the listener's visual perspective and the typicality of event components shape 4- and 5-year-olds' referential choices in production. We discuss our main findings and their significance below.

Interactive Contexts and Children's Referential Communication

A first key finding in our experiments was that children's informativeness was heavily context-dependent. In Experiment 1, where participants produced referential descriptions for the sake of a passive addressee, 4- and 5-year-olds often failed to provide the target disambiguating information (i.e., instruments, objects, or locations) that would allow their listener to identify the right event referent. By contrast, in Experiment 2, where participants engaged in a more genuine, collaborative interaction with a "true" interlocutor, both children's and adults' overall informativeness increased. Relatedly, in Experiment 1, 5-year-old children tended to mention more disambiguating information than 4-year-olds, but the difference disappeared in Experiment 2, where speakers' overall informativeness increased.

The role of context has important implications for theories about children's difficulties with reference production and pragmatic reasoning more generally. Recall that, for many researchers, children's problems with reference have to do with their difficulty understanding that referring means finding differences between a target referent and other objects in view (see Deutsch & Pechmann, 1982; Matthews et al., 2012, 2007; Nilsen & Mangal, 2012; Whitehurst & Sonnenschein, 1981). This position predicts that if the contrast between the target and its alternatives is made sufficiently salient, children should be likely to provide informative referential descriptions. Our data do not support this view: Although identifying the differences between potential referents is a prerequisite for successful disambiguation (but see Davies & Kreysa, 2018), the results of Experiment 1 suggest that presenting stimuli as closely matched pairs of contrasting events (and, thus, making the differences between the two potential referents highly noticeable) was not enough for eliciting informative descriptions.

Our results suggest that an important component of children's successes and failures with reference identification lies with the communicative circumstances of different experimental tasks. As Experiment 2 clearly shows, 4- and 5-year-olds (and adults) in our sample were more likely to be informative when communicating with an active listener in a task that required coordination of communicative efforts. This finding is consistent with—and can offer a framework for explaining—several prior findings from referential communication tasks. In prior work, children showed sensitivity to their communicative partner's needs when this partner was a real person who needed help to complete a clearly defined task, such as construct a toy (e.g., Bahtiyar & Küntay, 2009), move an object on a visual array (e.g., Nadig & Sedivy, 2002) or find a hidden toy (e.g., O'Neill, 1996). By contrast, children frequently produced underinformative utterances when asked to communicate with listeners who were either unable to react (e.g., because they were imaginary; see Girbau, 2001) or had no real stakes in the interaction (e.g., because they were the primary experimenter; see Davies & Kreysa, 2018; Rabagliati & Robertson, 2017).

How exactly did the interactive exchange in Experiment 2 affect informativeness? One possibility is that the more explicit goal and detailed exchanges between the active addressee and the speaker in Experiment 2 offered numerous and reliable cues as to the communicative purpose of the task. Before each trial, the addressee explicitly invited participants to identify a single referent ("I can see two pictures. Which one is it? Tell me about it!"). After each trial, the addressee had to guess the right picture on the basis of the speaker's verbal description and place a sticker on it; the addressee reacted to these demands of the guessing game with some uncertainty ("I hope I got it right!"). Furthermore, during the introductory phase, participants were given feedback after offering an ambiguous utterance and gained first-hand access to the addressee's visual perspective (e.g., children sat on the addressee's chair to make sure she could/could not see the computer screen). All of these features of the task may have offered explicit cues about the need for clear, unambiguous event descriptions, unlike Experiment 1 where the same requirements remained implicit and had to be inferred.[3] Other work has shown that young children, unless given

[3] In particular, corrective feedback is an effective way to increase informativeness in children's speech (e.g., see Bacso & Nilsen, 2017; Deutsch & Pechmann, 1982; Matthews et al., 2012, 2007; Nilsen & Mangal, 2012; Uzundag & Küntay, 2018; Wardlow & Heyman, 2016). However, it is unclear whether the feedback provided in Experiment 2 (i.e., placing the sticker next to the incorrect picture following an ambiguous utterance) was more salient than the feedback provided in Experiment 1 (i.e., pointing out the differences between the target and the distractor picture), as both types of feedback have been proven highly effective.

Table 8
Parameter Estimates for Length of Utterance Across Experiments 1–2

Effects	Estimate	*SE*	*t*
Intercept	6.89	.23	21.86***
Age (Adults vs. Children)	−2.89	.40	−7.29***
Age (5- vs. 4-year-olds)	.15	.45	.33
Experiment (1 vs. 2)	2.1	.37	5.66***
Age (Adults vs. Children): Experiment (1 vs. 2)	3.82	.80	4.81***
Age (5- vs. 4-year-olds): Experiment (1 vs. 2)	2.20	.89	2.48*

* $p < .05$. *** $p < .001$.

very explicit cues, fail to meet referential demands in production (e.g., Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Matthews, Butcher, Lieven, & Tomasello, 2012) and have inconsistent expectations about how informative speakers should be in comprehension (e.g., Katsos & Bishop, 2011; Papafragou & Musolino, 2003; Pouscoulous, Noveck, Politzer, & Bastide, 2007). Nevertheless, in our data, this line of reasoning faces the problem that the increase in 4- and 5-year-olds' informativeness between Experiments 1 and 2 was not restricted to disambiguating instruments, objects, and locations (e.g., mention of agents also increased even though this component was not task-relevant) and did not occur consistently for the sake of an addressee with limited knowledge.

A second, more likely possibility is that the collaborative context in Experiment 2 increased participants' motivation for communication by maximizing speakers' expected gains and offsetting the costs of designing fully informative utterances (on the cost-gains balance in communication, see Sperber & Wilson, 1986/1995; cf. also Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). Prior evidence has showed that 6- to 7-year-olds (but not 3- to 4-year-olds) provided more informative utterances in a referential production task when they were offered tangible incentives (stickers) for successful descriptions (Varghese & Nilsen, 2013; see also Bahtiyar & Küntay, 2009, Exp. 2). Although we did not provide such concrete incentives here, the collaborative context of Experiment 2 (similarly to other collaborative exchanges in real life) may have been inherently motivational for participants. As any other joint action, collaborative communication yields mutual benefits for participants by allowing them to gain access to resources that would otherwise be inaccessible (Melis & Warneken, 2016). Depending on the purpose of the exchange, the benefits can be straightforward (e.g., acquire information, get assistance) or psychologically more complex (e.g., express one's feelings, commit to a future behavior, etc.). In our studies, speakers may have been more likely to assist the collaborative interlocutor because of the shared psychological benefits of "winning" in the game or even "helping" someone else (cf. Hamann, Warneken, & Tomasello, 2012). Relatedly, given that referential communication can impose considerable cognitive costs for both children and adults (e.g., Brown-Schmidt, 2009; Epley, Morewedge, & Keysar, 2004; Lin, Keysar, & Epley, 2010; Nilsen & Graham, 2009; Nilsen et al., 2015; Roßnagel, 2000), speakers in our study may have been more likely to assume these costs when communicating with a collaborative interlocutor who reciprocally shared some of the burdens (by talking more, as in Experiment 2) instead of a passive listener who only enjoyed the benefits of the interaction without paying any of the costs (by talking very little, as in Experiment 1).

The precise explanation of the role of interactive context impacts the account of the change in children's performance between the ages of 4 and 5 (recall that, in Experiment 1, 5-year-olds tended to mention more disambiguating information than 4-year-olds but this difference disappeared in Experiment 2). One possibility is that what children lack compared with adults is the ability to spontaneously evaluate how much information is needed (a component of communicative competence that requires experience in social interaction; see Nilsen & Fecica, 2011). If so, younger children in our sample, as less experienced communicators, may have relied more heavily on the presence of relevant cues: Four-year-olds were less informative compared with 5-year-olds in the absence of such cues (in Experiment 1) but became as informative as 5-year-olds when these cues were abundant (in Experiment 2). This possibility is in line with other work showing that 4- and 5-year-old comprehenders do not penalize underinformativeness in the absence of relevant cues but even 4-year-olds improve when such cues are provided (Gweon & Asaba, 2017; Skordos & Papafragou, 2016). Furthermore, this work also demonstrated that the ability to entertain appropriate expectations of informativeness in the absence of relevant cues was the source of developmental change between preschoolers and school-age children (Gweon & Asaba, 2018). Interestingly though, these studies, unlike our own, did not find developmental differences between 4- and 5-year-olds.

Alternatively, what changes in development may be the motivation to communicate when the interaction is not particularly rewarding (or the balance between costs and benefits is not favorable). If so, younger children in our experiments may have relied more heavily on explicit (social) rewards (e.g., winning in a game) compared with older children: In the absence of such rewards (Experiment 1), 4-year-olds' referential communication lagged behind that of 5-year-olds. A related possibility is that younger children in our experiments were more likely to bear the conversation costs and offer the required information when communicating with a collaborative addressee, not only to increase their own benefits (win in the game) but perhaps also to increase their partner's benefits (help her find the right picture). In support of this possibility, Experiment 2 (control items) was the only case where our youngest sample displayed sensitivity to the listener's visual access—presumably a costly adaptation that requires sophisticated mentalizing skills (see also next section). This possibility resonates with recent findings showing that 4-year-olds—but not 5-year-olds—have limitations estimating the tradeoff between informativeness and communicative efficiency both in their own and in others' nonlinguistic demonstrations of facts (Gweon, Shafto, & Schulz, 2018).

Because Experiment 2 introduced several modifications to the paradigm of Experiment 1, the present data cannot definitively adjudicate between the two possibilities and the subsequent accounts of the mechanisms that guide children's communicative development. Additional research will be required to identify whether it is the presence of relevant cues or the motivation provided by collaborative communication that determines children's decisions of what to include in their linguistic messages.[4] Future research could tease these two possibilities apart by disentangling the role of cues to a task's communicative goal from participants' motivation to fulfill that goal collaboratively. Furthermore, if the ability to effectively estimate the tradeoff between costs and benefits is what defines children's informativeness in production, it remains to be seen whether these estimates concern children's own (egocentric) costs and benefits or those of their listener's as well.

Production Adaptations in Children's Event Reference

A second, striking aspect of our data is that—despite sensitivity to the conversational context—4- and 5-year-olds' productions often failed to address their communicative partner's needs. Even in Experiment 2, where informativeness increased, children's mention of disambiguating information remained nonadult-like: Five-year-olds mentioned instruments 55% of the time and disambiguating features 56% of the time, whereas 4-year-olds did so 43% and 47% of the time, respectively. This is not unprecedented. In many well-known referential communication studies, 5-year-olds appropriately modified their descriptions to uniquely identify the target referent less than half of the time (e.g., 30% for 5-year-olds in Bahtiyar & Küntay, 2009, Exp. 1; 39% for 5-year-olds in Nilsen & Graham, 2009; 45% in Davies & Katsos, 2010). Only studies that provided referential feedback report high levels of informativeness in children's repairs (e.g., see Deutsch & Pechmann, 1982; Matthews et al., 2012, 2007).

Two further sets of key findings throw light on the specific difficulties children face in their production adaptations. Both preschoolers and adults in our sample made adjustments to the typicality of events by mentioning atypical instruments more frequently than typical instruments across experiments. This result is in accordance with work on adult production showing that, in storytelling, atypical (i.e., highly unpredictable) instruments are more noteworthy than typical (i.e., inferable) instruments (Brown & Dell, 1987; Lockridge & Brennan, 2002), and recent developmental findings showing that children tend to mention unpredictable adjective-noun combinations (e.g., "cuddly pajamas") when these carry high information content (Bannard, Rosner, & Matthews, 2017; cf. also Papafragou et al., 2006). However, participants did not consistently adjust to the listener's visual access. Adults were highly informative in both visual access conditions, possibly because they realized that disambiguating information should be provided independently of whether the listener could see the target events or not. Children did not adjust their utterances to their listener's visual access (despite being often underinformative) but occasionally incorporated the listener's visual perspective: Five-year-olds in both experiments mentioned agents more frequently when the listener did not have visual access to the target events (compared with when she had access to them), and 4-year-olds in Experiment 2 mentioned disambiguating objects and locations more frequently when the listener could not see the events.

The current pattern of results can be explained by assuming that different types of adjustments impose distinct cognitive demands. Typicality based adjustments might have been easier to make because these required children to only consider what is usual within the community (or even what they themselves consider noteworthy) without any need to monitor the beliefs of a particular interlocutor. By contrast, adjustments to a listener's visual access might have been harder because such adjustments required tracking what the listener knew at each point in the discourse—which in the no visual access condition conflicted with what the child could access—and thus presupposed a dynamic model of the conversation that needed to be updated frequently (see Arnold, 2008, for discussion). Adjustments to an interactive addressee would fall between the other two in terms of audience design demands because, although linked to a particular listener and context, they did not require complex computations about this particular person's mental states but only a coarse estimate of the listener's need for information. In support of this last point, even though 4- and 5-year-olds in our sample became overall more informative when communicating with an interactive partner in Experiment 2, they did not offer more disambiguating information when their partner lacked visual access to the events (as they should were they specifically monitoring that partner's mental state).

Together these findings show that 4- and 5-year-old children do not consistently adjust their informativeness levels to the needs of their specific interlocutor and often direct their production toward what is noteworthy on a more general level (see also Bannard et al., 2017; Brown & Dell, 1987, for similar conclusions). More broadly, these observations highlight the need for a nuanced model of children's audience design, where some listener needs might be costlier to incorporate than others.

The present experiments went beyond prior work on nominal reference to focus on how children describe and identify events. Nevertheless, our investigation still used a disambiguation paradigm with a restricted conversational goal. Future work should extend the present results to tasks with more open-ended goals, such as situations in which children describe individual events freely for a partner. It remains to be seen whether the communicative status of that partner (e.g., whether he or she shares a joint goal with the child) might affect mention of instruments

[4] Notice that the factors shaping children's informativeness in production may be somewhat different from those in comprehension. In comprehension, preschoolers' sensitivity to informativeness seems to depend on the presence of relevant cues to the goals of the exchange (Gweon & Asaba, 2017; Skordos & Papafragou, 2016). However, in production, a process inherently constrained by the need of least effort for the speaker (Shannon, 1948; Zipf, 1949), the informativeness of children's messages may be more heavily affected by the motivation for communication (including considerations of the balance between costs and benefits).

We thank the children and the parents in the preschools of Newark, DE and Hockessin, DE area. We thank Elizabeth Abramson, U'nek Brown, Kendra Goodwin, Matt Cohen, Carlyn Friedberg, Sarah Lorch, Taraneh Mojaverian, Jillian Ferrara, CarolAnn Edie, Sarah Fairchild, Dan Curtin, and Hanna Schwartz for their assistance in these studies. Part of these data has been presented at TiXPrag1, CogSci, 2016 and BUCLD 41. This material is based upon work supported by the National Science Foundation under Grant 1632849.

and other event components. It also remains to be seen whether, under those circumstances, children might be more likely to consult the listener's visual perspective. We are currently pursuing these possibilities in ongoing work.

References

Adams, A.-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders, 35,* 95–116. http://dx.doi.org/10.1080/136828200247278

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes, 23,* 495–527. http://dx.doi.org/10.1080/01690960801920099

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511801686

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Bacso, S. A., & Nilsen, E. S. (2017). What's that you're saying? Children with better executive functioning produce and repair communication more effectively. *Journal of Cognition and Development, 18,* 441–464. http://dx.doi.org/10.1080/15248372.2017.1336438

Bahtiyar, S., & Küntay, A. C. (2009). Integration of communicative partner's visual perspective in patterns of referential requests. *Journal of Child Language, 36,* 529–555. http://dx.doi.org/10.1017/S0305000908009094

Bannard, C., Rosner, M., & Matthews, D. (2017). What's worth talking about? Information theory reveals how children balance informativeness and ease of production. *Psychological Science, 28,* 954–966. http://dx.doi.org/10.1177/0956797617699848

Barr, D. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59,* 457–474. http://dx.doi.org/10.1016/j.jml.2007.09.002

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1482–1493. http://dx.doi.org/10.1037/0278-7393.22.6.1482

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation, 53,* 301–344. http://dx.doi.org/10.1016/S0079-7421(10)53008-1

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology, 19,* 441–472. http://dx.doi.org/10.1016/0010-0285(87)90015-6

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review, 16,* 893–900. http://dx.doi.org/10.3758/PBR.16.5.893

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning? *Journal of Memory and Language, 54,* 592–609. http://dx.doi.org/10.1016/j.jml.2005.12.008

Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2016). How children and adults encode causative events cross-linguistically: Implications for language production and attention. *Language, Cognition and Neuroscience, 31,* 1015–1037. http://dx.doi.org/10.1080/23273798.2016.1175649

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition, 122,* 135–149. http://dx.doi.org/10.1016/j.cognition.2011.10.002

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software, 80,* 1–28. http://dx.doi.org/10.18637/jss.v080.i01

Davies, C. N., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua, 120,* 1956–1972. http://dx.doi.org/10.1016/j.lingua.2010.02.005

Davies, C., & Kreysa, H. (2017). Looking at a contrast object before speaking boosts referential informativeness, but is not essential. *Acta Psychologica, 178,* 87–99. http://dx.doi.org/10.1016/j.actpsy.2017.06.001

Davies, C., & Kreysa, H. (2018). Look before you speak: Children's integration of visual information into informative referring expressions. *Journal of Child Language, 45,* 1116–1143. http://dx.doi.org/10.1017/S0305000918000120

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition, 11,* 159–184. http://dx.doi.org/10.1016/0010-0277(82)90024-5

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40,* 760–768. http://dx.doi.org/10.1016/j.jesp.2004.02.002

Ford, W., & Olson, D. (1975). The elaboration of the noun phrase in children's description of objects. *Journal of Experimental Child Psychology, 19,* 371–382. http://dx.doi.org/10.1016/0022-0965(75)90068-5

Girbau, D. (2001). Children's referential communication failure: The ambiguity and abbreviation of message. *Journal of Language and Social Psychology, 20,* 81–89. http://dx.doi.org/10.1177/0261927X01020001004

Golinkoff, R. M. (1986). 'I beg your pardon?': The preverbal negotiation of failed messages. *Journal of Child Language, 13,* 455–476. http://dx.doi.org/10.1017/S0305000900006826

Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). What's learned together stays together: Speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 843–853. http://dx.doi.org/10.1037/a0029467

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). New York, NY: Academic Press.

Gweon, H., & Asaba, M. (2017). Order matters: Children's evaluation of underinformative teachers depends on context. *Child Development, 89,* e278–e292. http://dx.doi.org/10.1111/cdev.12825

Gweon, H., & Asaba, M. (2018). Order matters: Children's evaluation of underinformative teachers depends on context. *Child Development, 89,* e278–e292. http://dx.doi.org/10.1111/cdev.12825

Gweon, H., Shafto, P., & Schulz, L. (2018). Development of children's sensitivity to overinformativeness in learning and teaching. *Developmental Psychology, 54,* 2113–2125. http://dx.doi.org/10.1037/dev0000580

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development, 83,* 137–145. http://dx.doi.org/10.1111/j.1467-8624.2011.01695.x

Hartshorne, J. K., Pogue, A., & Snedeker, J. (2015). Love is hard to understand: The relationship between transitivity and caused events in the acquisition of emotion verbs. *Journal of Child Language, 42,* 467–504. http://dx.doi.org/10.1017/S0305000914000178

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59,* 91–117. http://dx.doi.org/10.1016/0010-0277(96)81418-1

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension.

Journal of Psycholinguistic Research, 29, 597–626. http://dx.doi.org/10.1023/A:1026468209238

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences, 20,* 589–604. http://dx.doi.org/10.1016/j.tics.2016.05.011

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). United Kingdom: Oxford University Press.

Katsos, N., & Bishop, D. V. M. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition, 120,* 67–81. http://dx.doi.org/10.1016/j.cognition.2011.02.015

Koenig, J.-P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition, 89,* 67–103. http://dx.doi.org/10.1016/S0010-0277(03)00082-9

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes, 47,* 567–587. http://dx.doi.org/10.1080/01638530903441339

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review, 20,* 54–72. http://dx.doi.org/10.3758/s13423-012-0341-8

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science, 36,* 517–544. http://dx.doi.org/10.1111/j.1551-6709.2011.01220.x

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology, 46,* 551–556. http://dx.doi.org/10.1016/j.jesp.2009.12.019

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review, 9,* 550–557. http://dx.doi.org/10.3758/BF03196312

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: A review of associations with formal language, social cognition, and executive functions. *Language Learning and Development, 14,* 186–223. http://dx.doi.org/10.1080/15475441.2018.1455584

Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two- and four-year-olds learn to adapt referring expressions to context: Effects of distracters and feedback on referential communication. *Topics in Cognitive Science, 4,* 184–210. http://dx.doi.org/10.1111/j.1756-8765.2012.01181.x

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions. *Applied Psycholinguistics, 27,* 403–422. http://dx.doi.org/10.1017/S0142716406060334

Matthews, D., Lieven, E., & Tomasello, M. (2007). How toddlers and preschoolers learn to uniquely identify referents for others: A training study. *Child Development, 78,* 1744–1759. http://dx.doi.org/10.1111/j.1467-8624.2007.01098.x

McDaniel, D., McKee, C., & Garrett, M. F. (2010). Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language, 37,* 59–94. http://dx.doi.org/10.1017/S0305000909009507

Melis, A. P., & Warneken, F. (2016). The psychology of cooperation: Insights from chimpanzees and children. *Evolutionary Anthropology, 25,* 297–305. http://dx.doi.org/10.1002/evan.21507

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science, 13,* 329–336. http://dx.doi.org/10.1111/j.0956-7976.2002.00460.x

Nilsen, E. S., & Fecica, A. M. (2011). A model of communicative perspective-taking for typical and atypical populations of children. *Developmental Review, 31,* 55–78. http://dx.doi.org/10.1016/j.dr.2011.07.001

Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology, 58,* 220–249. http://dx.doi.org/10.1016/j.cogpsych.2008.07.002

Nilsen, E. S., & Graham, S. A. (2012). The development of preschoolers' appreciation of communicative ambiguity. *Child Development, 83,* 1400–1415. http://dx.doi.org/10.1111/j.1467-8624.2012.01762.x

Nilsen, E. S., & Mangal, L. (2012). Which is important for preschoolers' production and repair of statements: What the listener knows or what the listener says? *Journal of Child Language, 39,* 1121–1134. http://dx.doi.org/10.1017/S0305000911000432

Nilsen, E. S., Varghese, A., Xu, Z., & Fecica, A. (2015). Children with stronger executive functioning and fewer ADHD traits produce more effective referential statements. *Cognitive Development, 36,* 68–82. http://dx.doi.org/10.1016/j.cogdev.2015.09.001

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development, 67,* 659–677. http://dx.doi.org/10.2307/1131839

O'Neill, D. K., & Topolovec, J. C. (2001). Two-year-old children's sensitivity to the referential (in)efficacy of their own pointing gestures. *Journal of Child Language, 28,* 1–28. http://dx.doi.org/10.1017/S0305000900004566

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition, 98,* B75–B87. http://dx.doi.org/10.1016/j.cognition.2005.05.005

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition, 86,* 253–282. http://dx.doi.org/10.1016/S0010-0277(02)00179-8

Perner, J., & Leekam, S. R. (1986). Belief and quantity: Three-year-olds' adaptation to listener's knowledge. *Journal of Child Language, 13,* 305–315. http://dx.doi.org/10.1017/S0305000900008072

Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition, 14,* 347–375. http://dx.doi.org/10.1080/10489220701600457

Rabagliati, H., & Robertson, A. (2017). How do children learn to avoid referential ambiguity? Insights from eye-tracking. *Journal of Memory and Language, 94,* 15–27. http://dx.doi.org/10.1016/j.jml.2016.09.007

Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.), *Sociological methodology 1995* (pp. 111–196). Cambridge, MA: Blackwell.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Rissman, L., Rawlins, K., & Landau, B. (2015). Using instruments to understand argument structure: Evidence for gradient representation. *Cognition, 142,* 266–290. http://dx.doi.org/10.1016/j.cognition.2015.05.015

Roßnagel, C. (2000). Cognitive load and perspective-taking: Applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology, 30,* 429–445. http://dx.doi.org/10.1002/(SICI)1099-0992(200005/06)30:3<429::AID-EJSP3>3.0.CO;2-V

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27,* 379–423. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition, 153,* 6–18. http://dx.doi.org/10.1016/j.cognition.2016.04.006

Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication: A hierarchy of skills. *Child Development, 55,* 1936–1945. http://dx.doi.org/10.2307/1129940

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition* (2nd ed. 1995). Cambridge, MA: Harvard University Press.

Uzundag, B. A., & Küntay, A. C. (2018). Children's referential communication skills: The role of cognitive abilities and adult models of speech. *Journal of Experimental Child Psychology, 172,* 73–95. http://dx.doi.org/10.1016/j.jecp.2018.02.009

Varghese, A. L., & Nilsen, E. (2013). Incentives improve the clarity of school-age children's referential statements. *Cognitive Development, 28,* 364–373. http://dx.doi.org/10.1016/j.cogdev.2013.07.001

Wardlow, L., & Heyman, G. D. (2016). The roles of feedback and working memory in children's reference production. *Journal of Experimental Child Psychology, 150,* 180–193. http://dx.doi.org/10.1016/j.jecp.2016.05.016

Whitehurst, G. J. (1976). The development of communication: Changes with age and modeling. *Child Development, 47,* 473. http://dx.doi.org/10.2307/1128804

Whitehurst, G. J., & Sonnenschein, S. (1981). The development of informative messages in referential communication: Knowing when versus knowing how. In W. I. Dickson (Ed.), *Children's oral communication skills* (pp. 127–142). New York, NY: Academic Press.

Wilson, F., Papafragou, A., Bunger, A., & Trueswell, J. (2011). Rapid extraction of event participants in caused motion events. In L. A. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Proceedings of the 33th Annual Conference of the Cognitive Science Society* (pp. 1206–1211). Austin, TX: Cognitive Science Society.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison Wesley Press.

Appendix A

Test Stimuli

	Disambiguating instrument	
Event	Atypical	Typical
1. A boy is cutting paper	using scissors	using a knife
2. A woman is cleaning	using a cloth	using a teddy bear
3. A man is eating a steak	using a fork	using a knife
4. A man is holding a door open	using his hand	using a stick
5. A girl is coloring a picture	using a red crayon	using a red lipstick
6. A woman is sweeping the floor	using a broom	using a branch
7. A woman is blowing her nose	using a tissue	using a shirt
8. A woman is brushing her teeth	using a toothbrush	using her finger

Appendix B

Control Stimuli

Event	Disambiguating object
1. A woman is holding	A bottle (vs. a spoon)
2. A baby is holding	A rattle (vs. a kite)
3. A woman is putting	A ball (vs. a star) onto a Christmas tree
4. A woman is spreading	A newspaper (vs. red paper) onto a table
	Disambiguating location
1. A man is fixing a fence	Next to a tree (vs. no tree)
2. A man is watering plants	Next to a bench (vs. no bench)
3. A couple is sewing a blanket	Under a tent (vs. no tent)
4. A man is stretching	Near a fridge (vs. no fridge)

Note. Nontarget items appear in parentheses.

Received July 26, 2017
Revision received November 26, 2018
Accepted December 12, 2018 ■