COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12790

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12790

Children's (and Adults') Production Adjustments to Generic and Particular Listener Needs

Myrto Grigoroglou,[a] Anna Papafragou[b,*]

[a]*Department of Linguistics and Cognitive Science, University of Delaware*
[b]*Department of Psychological and Brain Sciences, University of Delaware*

Received 27 December 2018; received in revised form 21 August 2019; accepted 26 August 2019

Abstract

Adults design utterances to match listeners' informational needs by making both "generic" adjustments (e.g., mentioning atypical more often than typical information) and "particular" adjustments tailored to their specific interlocutor (e.g., including things that their addressee cannot see). For children, however, relevant evidence is mixed. Three experiments investigated how generic and particular factors affect children's production. In Experiment 1, 4- to 5-year-old children and adults described typical and atypical instrument events to a silent listener who could either see or not see the events. In later extensions, participants described the same events to either a silent (Experiment 2) or an interactive (Experiment 3) addressee with a specific goal. Both adults and 4- to 5-year-olds performed generic adjustments but, unlike adults, children made listener-particular adjustments inconsistently. These and prior findings can be explained by assuming that particular adjustments can be costlier for children to implement compared to generic adjustments.

Keywords: Typicality; Event cognition; Common ground; Language production; Instruments; Informativeness

1. Introduction

How do people decide what to say when communicating with others? According to traditional views of communication, speakers strive to make contributions that are informative and relevant to the purpose of their communicative exchanges with others (Grice, 1975) by taking into account *common ground*, the set of knowledge and beliefs they share with a communicative partner via a process often referred to as *audience design*

Correspondence should be sent to Myrto Grigoroglou, Department of Applied Psychology and Human Development, University of Toronto, 252 Bloor Street West, Toronto ON M5S 1V6, Canada. E-mail: myrto31@gmail.com

*Present address: Department of Linguistics, University of Pennsylvania, Philadelphia, PA, 19104,USA

(Clark & Marshall, 1981). Common ground can be established at various levels of generality and includes information interlocutors share in their immediate perceptual context (as physically co-present in the same environment), in prior discourse (as participants in the same conversation) or within the community (as members of the same group; Clark & Marshall, 1981). Experimental evidence suggests that adults can adjust their speech to their listeners' informational needs by using various sources of shared information, such as visual co-presence with a listener (e.g., Horton & Keysar, 1996; Lockridge & Brennan, 2002; Nadig & Sedivy, 2002; Yoon, Koh, & Brown-Schmidt, 2012), common experience in prior discourse (e.g., Brennan & Clark, 1996; Galati & Brennan, 2010; Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Heller, Gorman, & Tanenhaus, 2012; Horton & Gerrig, 2002, 2005; Wu & Keysar, 2007), or knowledge generally available within a linguistic or cultural community (e.g., Clark & Marshall, 1981; Uther, Knoll, & Burnham, 2007).

For children, however, evidence for adjustments to listener needs includes contrasting findings. Prior research has mostly focused on children's ability to make adjustments to the needs of a listener for the purposes of referential disambiguation (i.e., when children need to uniquely specify a target object among similar-looking distractors). Some studies in this line of work have shown that children, even at a very young age, can successfully adjust the informational content of their pointing gestures or simple utterances to the visual perspective of their listener (e.g., Bahtiyar & Küntay, 2009; Nadig & Sedivy, 2002; Nilsen & Graham, 2009; O'Neill, 1996). However, other studies reveal rampant failures to include disambiguating information in children's speech: In many cases, preschoolers successfully adapt their referential productions to their partner's needs in less than half of the critical trials (e.g., Bahtiyar & Küntay, 2009; Davies & Katsos, 2010; Nilsen & Graham, 2009) or ignore their listener's perspective altogether (e.g., Bannard, Rosner, & Matthews, 2017; Girbau, 2001) and continue to offer underinformative descriptions even until late childhood (e.g., Deutsch & Pechmann, 1982; Fukumura, 2016; Sonnenschein & Whitehurst, 1984).

A separate, more limited set of studies has looked at how children describe events to their interlocutors. Events involve multiple participants and the relations between them, and unfold in time and space. When describing an event, speakers have great flexibility in terms of which aspect of an event to comment on, and a single, observable event can be described at multiple levels of specificity ("A boy and a girl are running from a beach house towards the sea carrying a dragon-shaped kite at sunset," "Two children are playing on a beach," and "Summer fun," can all be used to describe the same scene). Available evidence, similarly to the disambiguation literature, has shown that children have the ability to adjust their event descriptions to what their listeners are likely to know but frequently omit (what to adults are) critical pieces of information. For instance, in one study, 3-year-olds successfully described an event to an addressee who had not witnessed it; nevertheless, younger 3-year-olds were likely to mention only one of two events, even when their addressee had seen neither (Perner & Leekam, 1986; see also Allen, Hughes, & Skarabela, 2015). Children have also been found to produce sparse, non-adult event descriptions when shown dynamic events and asked to say "what happened." In one demonstration (Bunger, Trueswell, & Papafragou, 2012), adults predominantly described both the manner and the path of a motion event ("The boy was skating into the net"), whereas 4-year-olds were more likely to mention only the manner ("The boy was skating"). Similarly, in a question-answering task, 2- to 4-year-old children gave overwhelmingly underinformative responses when asked to say "what's happening" in an event (although they had no difficulty answering more specific questions, for example, "What's X doing?"; Salomo, Lieven, & Tomasello, 2013). In sum, the literature points to both early successes and important limitations in children's ability to produce utterances that are informative for listeners.

1.1. *Explanations of children's production choices*

Several different explanations for children's underinformative production have been proposed. One possibility is that children omit information needed by addressees because, unlike adults, they fail to direct their attention to critical scene components that should find their way into grammatical encoding (for discussion on the relation between scene apprehension and linguistic production in children, see Bunger et al., 2012; Davies & Kreysa, 2018; Deutsch & Pechmann, 1982; Pechmann, 1989; Whitehurst, 1976; Whitehurst & Sonnenschein, 1981). Despite the plausibility of this explanation, recent eye-tracking studies with children have shown that eye-gaze patterns do not always predict which scene component will be mentioned in production (e.g., Bunger et al., 2012; Davies & Kreysa, 2018; cf. Rabagliati & Robertson, 2017). For instance, in a previously mentioned motion event description study (Bunger et al., 2012), eye tracking evidence showed that attention allocation to manner and path regions of the motion events did not differ between children and adults, despite the fact that adults mentioned both event components but children only one. Thus, children's linguistic information omissions do not always seem to be due to problems with scene apprehension per se.

Alternatively, children may face limitations in linguistic or cognitive resources. It is possible that children have limited linguistic capacity during utterance preparation (e.g., see Adams & Gathercole, 2000; Levelt, 1989; McDaniel, McKee, & Garrett, 2010), especially when they describe complex visual stimuli such as events. Describing events is more complex than describing objects since events may have multiple participants, each of whom might be described at various levels of detail. Communicating "who did what to whom," although essential in everyday life, might pose particular challenges as it requires one to construct a cognitive representation of the event but also to place event participants (e.g., agents, patients, instruments) into specific syntactic positions (e.g., subject, direct object, adjunct phrase). Yet in many cases, it appears that children fail to produce adequately informative utterances even at an age when they have acquired the necessary lexical and syntactic resources (e.g., path modifiers in Bunger et al., 2012).

Furthermore, children may produce utterances that are underinformative for listeners because of limitations in cognitive abilities such as working memory and inhibition, or mentalizing skills such as Theory of Mind (Allen, Skarabela, & Hughes, 2008; Bannard et al., 2017; Brown-Schmidt, 2009; Epley, Morewedge, & Keysar, 2004; Nilsen & Fecica, 2011; Nilsen & Graham, 2009; Nilsen, Varghese, Xu, & Fecica, 2015; Wardlow & Heyman, 2016). In support of this possibility, informative language use in children during reference disambiguation tasks is associated with increased ability to explicitly report what each conversational partner knows (Roberts & Patterson, 1983), better performance on independent Theory of Mind tasks (Resches & Pérez Pereira, 2007), and stronger working memory (e.g., Nilsen & Graham, 2009; Nilsen et al., 2015; Wardlow & Heyman, 2016). At present, however, it is not clear whether standardized, global measures of cognitive skills are the best predictor of the entire range of children's pragmatic abilities or whether different types of adjustments to listeners engage specific cognitive mechanisms to different degrees (see also Roberts & Patterson, 1983; Wardlow & Heyman, 2016).

A less studied possibility is that children's difficulty may lie with pragmatic decisions about "what is worth talking about" from the perspective of a listener, which may go beyond considerations of the listener's unique knowledge. These decisions rely on a variety of factors such as the need to comment on things that are unexpected (Greenfield, 1980), to use locutions that are conventional (Clark, 2007), to deviate from conventional choice of words to mark a less stereotypical status quo (Levinson, 1995), and to keep utterances short to reduce production costs for the speaker and processing costs for the listener (Zipf, 1949). Such pragmatic decisions interact dynamically with speakers' non-linguistic conceptualization of the message they want to convey and affect the way speakers package information into specific syntactic units (Levelt, 1989; Papafragou & Grigoroglou, 2019). Current models of audience design in children typically lack the nuance required to explain how children's non-optimal production choices may be affected by limitations in each of these components of the production process (but see Bannard et al., 2017).

1.2. A nuanced model of audience design

Recent pragmatic theories and adult psycholinguistic literature have suggested that not all adjustments in production are made by having a rich model of the listener's beliefs in mind, but some adjustments are based on less complex heuristics concerning what is relevant or salient in a specific situation. In pragmatic theory, this more nuanced position is captured by distinctions between pragmatic abilities that require mentalizing skills (e.g., Theory of Mind) and others that do not (e.g., see Levinson, 1995; O'Neill, 2012; Recanati, 2004; Sperber, 1994; see also Andrés-Roqueta & Katsos, 2017, for a review). Other theorists make more fine-grained distinctions based on the type and complexity of ToM skills required for different communication tasks (e.g., Sperber, 1994). Although viewed from the aspect of the comprehender rather than the speaker, these nuanced theoretical approaches to pragmatic meaning highlight the need to distinguish different types and levels of audience design.

In the adult psycholinguistics literature, Brown and Dell (1987; see also Dell & Brown, 1991) suggested that speakers make two broad types of adjustments in production. *Particular-listener adjustments* are adaptations made either to specific properties of the listener or to the context in which communication occurs. Adjusting one's speech to what a communicative partner can see or to information shared in prior discourse are characteristic examples of adjustments made for a particular listener. *Generic-listener adjustments* are adaptations made simply to facilitate comprehension by the average listener. For instance, adjustments to what is generally known within a linguistic or cultural community would fall under this category.

To explore this distinction, Brown and Dell (1987) had adults read and retell stories involving typical or atypical instruments (e.g., *Adolf stabbed the man with a knife* vs. *with an icepick*) to listeners who either had or did not have access to pictures depicting the events in the story. They found that speakers were more likely to mention atypical as opposed to typical instruments (a generic adaptation based on generally known facts), especially early in the speech planning process, within the same clause as the main verb (e.g., *Adolf stabbed the man with/using an ice pick*). However, the listener's knowledge did not affect the typicality effect: When the listener could not see the pictures, speakers were simply more likely to mention both kinds of instruments in separate clauses, after the verb (e.g., *Adolf stabbed the man. He used a knife/an ice pick*). They concluded that speakers' beliefs about addressees' knowledge do not affect utterance design early in the process of speech planning, but only later as a repair to an underinformative message (cf. Horton & Keysar, 1996; Keysar, Barr, & Horton, 1998; Pickering & Garrod, 2004). Later work by Lockridge and Brennan (2002) has challenged this conclusion: When they replicated the Brown and Dell (1987) study with naïve listeners instead of confederates, speakers were more likely to mention atypical instruments when the listener did not have access to the pictures depicting the events in the story, and did so early in speech planning (i.e., in the same clause as the action verb; see also Vanlangendonck, Willems, Menenti, & Hagoort, 2016, for similar evidence). Thus, speakers are more likely to make particular-listener adjustments when listeners have genuine informational needs (or their needs are *perceived* as genuine; see Buz, Tanenhaus, & Jaeger, 2016; Kuhlen & Brennan, 2010).

We propose that a flexible version of the distinction between particular- and generic-listener adjustments from the adult psycholinguistics literature (and similar proposals from the pragmatics literature) can be a powerful tool for developing a cognitive model for children's audience design that can organize and extend the currently disparate set of developmental findings. We take the position that children's successes and failures with different types of adjustments in production depend on whether these require a "particular" model of the listener's beliefs or a more "generic," less detailed model, as well as on the specific computational costs of building and maintaining these models (see also Hendriks, 2016; Moll & Kadipasaoglu, 2013; Moll & Tomasello, 2007). For example, adjusting one's speech to another person's visual perspective, especially if this changes on a regular basis and is different from the speaker's own perspective (as in standard referential disambiguation tasks) requires a computationally demanding "particular" model of the listener's beliefs that needs to be updated frequently. This type of particular-listener adaptation requires speakers to use their mentalizing skills, inhibit their own perspective, and hold a lot of information in their memory (see also Arnold, 2008). Unsurprisingly, this type of adaptation has elicited inconsistent findings in the developmental literature (e.g., Bannard et al., 2017; Matthews et al., 2006; Nadig & Sedivy, 2002).

At the other end, adjusting one's speech to information shared within a cultural or linguistic community only requires a "generic" model of the listener's beliefs (also shared by the speaker) without the need to consider what a specific partner knows, or update this set of beliefs during the interaction. Indeed, there is evidence that children tend to comment on unpredictable or atypical information—a "generic" adaptation. For example, in languages such as Greek where manner of motion is encoded selectively, both children and adults mention it when it is unexpected (e.g., when they see a man going up the stairs running) but not when it is predictable (e.g., when the man is going up the stairs walking; Papafragou, Massey, & Gleitman, 2006; cf. also Bannard et al., 2017; Köymen, Mammen, & Tomasello, 2016).

This analysis anticipates that, in many cases, generic-listener adjustments might be easier for children to implement than particular-listener adjustments because they are stable, consistent with the speaker's own perspective, and require less monitoring and updating. Furthermore, within particular-listener adjustments, computationally simpler adaptations (e.g., based on information shared in joint engagement with a particular interlocutor) would be easier for children and possibly emerge earlier than costlier adaptations (e.g., based on a particular listener's visual perspective, especially when that perspective conflicts with that of the child's). There is some evidence that these asymmetries might impact production adaptations even in adults: A closer look at Lockridge and Brennan's (2002) data shows that the effect of visual access was smaller than the effect of typicality.

The small number of developmental studies that have examined more than one type of adaptation supports our proposal. For instance, in an event description study where 2- to 4-year-olds were asked to describe the actions of a character in an event to a listener who could either see or not see this event, children's ability to make adjustments to information shared in prior discourse with the listener arose earlier than the ability to make adjustments to the listener's visual perspective (Matthews et al., 2006). In a referential disambiguation study where 4- and 5-year-olds were asked to uniquely identify a target event from a pair of almost identical events for a listener who could either see or could not see the events, children frequently made adjustments for a "generic" listener by mentioning generally unpredictable event information but did not make "particular" adjustments to the listener's visual perspective (Grigoroglou & Papafragou, 2019). And in another study, 3-year-olds' tendency to modify nouns with adjectives when describing pictures increased the more unpredictable the adjective-noun combinations became (e.g., *funny pajamas*), but this tendency was not affected by the listener's visual access to the pictures (Bannard et al., 2017). However, further research is needed on the development of the ability to make different kinds of listener adaptations because these studies either did not test adults (Bannard et al., 2017; Matthews et al., 2006) or included tasks that did not reveal adults' full sensitivity to the listener's visual perspective (Grigoroglou & Papafragou, 2019).

1.3. Current study

This study is the first to directly investigate the distinction between particular- and generic-listener adjustments in children's and adults' spontaneous event descriptions. Our paradigm was inspired by Brown and Dell (1987) and Lockridge and Brennan (2002) but, instead of having participants retell stories, we asked them to watch video-taped events and describe them to a listener. Spontaneous production, unlike story retelling, allows for examination of on-going speech planning processes as speakers encode unfolding perceptual information and make on-line syntactic choices. Precisely because of its immediacy, spontaneous speech, as opposed to retelling, may be more likely to induce particular-listener adaptations from speakers (see Galati & Brennan, 2010). Furthermore, the act of describing newly experienced events to other people is fundamental to everyday life, as it provides a way to organize the continuous flux of stimuli around us into meaningful units which can be shared and can scaffold later recollection of what happened (Zacks & Tversky, 2001).

In three experiments, we investigated the production adjustments made by 4- to 5-year-old children and adults. We chose this age group in children because of prior conflicting findings about audience design in this group. Similarly to Brown and Dell (1987) and Lockridge and Brennan (2002), we investigated *generic-listener* adjustments by manipulating instrument typicality in a series of everyday events (e.g., a man eating spaghetti with *a fork/a serving spoon*). We investigated *particular-listener* adjustments by manipulating the listener's visual access to the events. Similarly to these two studies, we expected that adults would be more likely to encode atypical as opposed to typical instruments (a *generic* adjustment) and that this tendency would be more pronounced when talking to addressees with limited visual access (a *particular* adjustment). Of interest was whether children would also perform these adaptations and to the same degree.

Our three experiments involved the same basic paradigm but differed in terms of the role and goals of the listener. Building on evidence that adults are more likely to make particular-listener adjustments when their listeners have (or are perceived to have) actual informational needs (Buz et al., 2016; Kuhlen & Brennan, 2010; Lockridge & Brennan, 2002) and developmental observations, suggesting that children are more likely to make production adjustments when the listener is an actual person with "real" informational needs (e.g., Grigoroglou & Papafragou, 2019; Köymen et al., 2016; Nadig & Sedivy, 2002) rather than an imaginary addressee in a more artificial exchange (e.g., Girbau, 2001), we hypothesized that speakers would become increasingly likely to offer instrument information, especially when the listener lacked visual access, as the listener's needs became more explicit across experiments.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants were 48 4- to 5-year-old children and 48 adults. The children were between the ages of 4;1 to 5;11, with a mean age of 5;0. All children were recruited from daycares in Newark, Delaware, and in a local children's museum and were tested at the location of recruitment. Adults were undergraduate students at the University of Delaware and received course credit for their participation. Informed consents was obtained by all participants (for children, consent was provided by their parents). The study was approved by the University of Delaware institutional review board and was carried out in accordance with the Declaration of Helsinki.

2.1.2. Materials

2.1.2.1. Pre-test: A separate group of 14 adults and 16 4-year-old children were given a preliminary task to determine instrument typicality for a set of 15 events. These groups were given a questionnaire (administered orally for the children) about the instruments used to perform certain everyday actions (i.e., "What do we use to eat spaghetti?", "What do we use to open a can?", etc.).

On the basis of participants' responses, we selected 12 events to be used in the main experiment, each with a typical and an atypical version. For the typical version, we selected an instrument that appeared frequently in both adults' and children's responses (e.g., a cloth for cleaning the counter). Overall, the selected typical instruments were mentioned in 73% of the adult responses and 63% of the child 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 (e.g., a stuffed animal for cleaning the counter).

2.1.2.2. Test materials: To depict each of the 12 test events, we created two short video clips, one showing the typical and the other the atypical version of the event (see Appendix A, for a full list). We also created a set of eight filler clips showing various everyday actions that did not involve instruments (e.g., reading a book, exercising). The same (male) agent performed all actions in the test and filler clips. The clips were arranged into two basic stimulus lists, each of which contained a single version of each test event and all filler events for a total of 20 stimuli. Each basic list contained six typical and six atypical instrument events.

2.1.3. Procedure

Participants were informed that they would be watching a set of short video clips and that, at the end of each video, they would have to describe what they saw. They were also introduced to a "friend" of the experimenter (an adult confederate listener) who had not seen the videos and wanted to know what participants would see. Participants were randomly assigned to one of two conditions. In the *Visual Access* condition, the listener sat next to the participant and also watched the clips. In the *No Visual Access* condition, the listener sat behind an opaque barrier so that she was unable to see the videos (or the participants as they were describing them). No restrictions were placed on participants' productions.

2.1.3.1. Coding: Responses from both adults and children were tape-recorded and later transcribed for coding purposes. Each response in the test items was coded for mention (or lack of mention) of the instrument in the video. We used Brown and Dell's (1987)

Table 1
Coding scheme and examples

A. Explicit mention	
Within Clause	
1.	After the verb He's nailing a nail into a fence *with a shoe*.
2.	Before the verb A guy using *his key* to get into the house.
3.	Incorporated into the verb The man is *hammering* a fence.
Separate Clause	
4.	After the verb A guy - I think he was hammering something, but he was using something that didn't look like a hammer . . . *almost like a shoe*.
5.	Before the verb A man getting *the knife* and cutting something.
B. Implicit mention	
6.	Atypical instrument inferable by the choice of locution. He is *cutting*. (opening a can with a knife) He is *folding* (a present). (wrapping a present with a towel) I see him bringing the *laundry* out. (catching a ball with a laundry basket) He is *hanging* something. (opening the door with a hanger)
C. Other	
7.	Not in the previous categories. I see *a plate*. (digging a hole with a plate)

coding scheme (also adopted by Lockridge & Brennan, 2002), which is summarized in Table 1. Briefly, we coded whether the instrument was mentioned *explicitly* (either within the same clause or in a separate clause), and whether it appeared before or after the main verb. We made one modification to Brown and Dell's scheme: We defined *implicit mention* of instruments as cases where the instrument was not mentioned but simply inferred.[1] This was especially relevant to certain responses in children's data, where the presence of an atypical instrument was inferable by the choice of locution (e.g., "He is cutting," for an event in which the agent was using a knife to open a can of food; "He is folding," for an event in which the agent was using a towel to wrap a present).[2] We also included very infrequent, *other* cases where the relevant object was mentioned but not in an instrument role. If speakers mentioned the instrument more than once, only the first mention was coded. To ensure inter-rater reliability, a random 20% sample of the data (across all three experiments) was coded by a second rater, blind to the visual access condition. Raters had 97% agreement (Cohen's $\kappa = 0.94$).

2.2. Results

2.2.1. Data analytic strategy

We analyzed the participants' descriptions by measuring (total) Instrument Mention. Because our task involved describing a single instrument event and children in our data

Table 2
Proportion of typical and atypical instrument mention for each age group and visual access condition in Experiment 1

	Children				Adults			
	Visual Access		No Visual Access		Visual Access		No Visual Access	
Category	Typ	Atyp	Typ	Atyp	Typ	Atyp	Typ	Atyp
Explicit mention								
Within Clause								
After the verb	0.01	0.14	0.01	0.22	0.14	0.66	0.27	0.72
Before the verb	0	0.04	0	0.03	0.05	0.09	0.07	0.17
Incorporated into the verb	0.06	0.02	0.06	0.02	0.05	0.02	0.07	0
Total	0.07	0.20	0.07	0.27	0.24	0.77	0.41	0.89
Separate Clause								
After the verb	0.01	0.02	0	0.01	0.01	0.05	0.02	0.02
Before the verb	0.01	0.05	0	0.01	0.04	0.05	0.04	0.06
Total	0.02	0.07	0	0.02	0.05	0.10	0.06	0.08
Implicit mention	0	0.05	0	0.05	0	0.01	0	0
Other	0.02	0.04	0.05	0.03	0	0	0	0
Total mention	0.11	0.36	0.12	0.37	0.29	0.88	0.47	0.97

rarely produced more than one sentence, we chose not to report within-clause and separate-clause Instrument mention analyses separately (but see Table 2 for the means for each category). Analyses of within-clause mentions across experiments simply replicated the findings from the analysis of total Instrument Mention.

The analysis dataset for Experiment 1 consisted of 96 subjects $\times$ 12 test items = 1,152 observations. Inspection of the data showed five missing cases (0.4%). Instrument Mention was a binary outcome variable coded as 1 (Instrument present) or 0 (Instrument absent). Data were analyzed using multi-level logistic mixed-effects modeling (Baayen, 2008; Baayen, Davidson, & Bates, 2008). Maximal random effect structure justified by the design was used in all models. In the event of convergence failure, random slopes were removed stepwise starting with the term with the least variance (Barr, Levy, Scheepers, & Tily, 2013). All models were fit using the *glmer* function of the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in the R Project for Statistical Computing (R Development Core Team, 2018).

2.2.2. Results

We analyzed the data using a model that included Mention of Instruments as the binary-dependent variable. Fig. 1 summarizes the data. The model included fixed effects of Typicality (Typical, Atypical) as a first-level fixed predictor, Age (Children, Adults) and Visual Access (Visual Access, No Visual Access) as second-level predictors, and all their interactions. The model also included random by-Participants intercepts and slopes for Typicality, as well as random by-Item intercepts and slopes for Typicality. The fixed effects of Typicality, Age and Visual Access were coded with centered contrasts (−0.5,

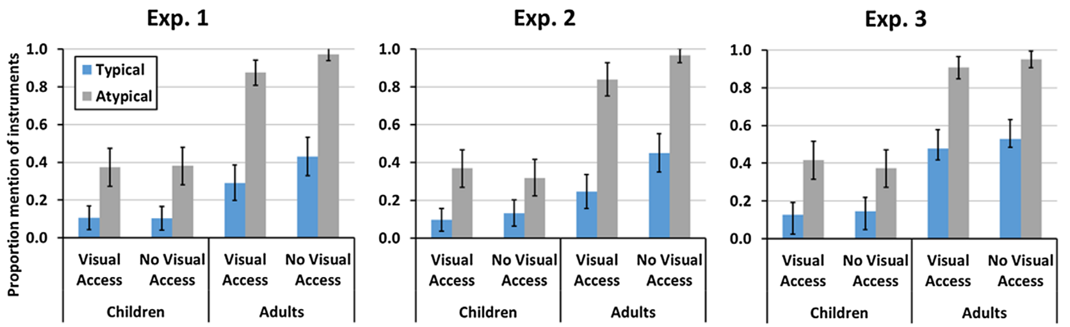

Fig. 1. Proportion of typical and atypical instrument mention per age group and visual access in Experiments 1–3. Error bars represent standard error.

0.5). The same coding strategy was followed in all subsequent analyses. Table 3 presents the parameter estimates for the multi-level model of Instrument Mention (see Appendix B for random effects). The model revealed a significant effect of Typicality: Atypical instruments were mentioned more often than typical instruments ($M_{Atypical} = 0.65$, $M_{Typical} = 0.24$). There was also a significant effect of Age: Adults mentioned instruments more frequently than children ($M_{Adults} = 0.65$, $M_{Children} = 0.24$). The model also returned a significant effect of Visual Access ($M_{Visual} = 0.41$, $M_{NoVisual} = 0.48$). These effects were qualified by two significant interactions between Typicality and Age and Age and Visual Access. We followed up these interactions by fitting two separate models in children and adults. Results showed that atypical instruments were used more frequently than typical instruments by both age groups but, because children mentioned instruments less, the magnitude of the difference was greater in adults ($\beta = 4.23$, $SE = 0.63$, $z = 6.67$, $p < .001$, $M_{Atypical} = 0.92$, $M_{Typical} = 0.38$) than in children ($\beta = 2.38$, $SE = 0.79$, $z = 3.02$, $p < .01$, $M_{Atypical} = 0.37$, $M_{Typical} = 0.11$). The interaction between Age and Visual Access was due to the fact that adults mentioned instruments more frequently when the listener could not see the events described compared to when she could see the events ($\beta = 1.48$, $SE = 0.44$, $z = 3.34$, $p < .001$, $M_{Visual} = 0.58$, $M_{NoVisual} = 0.72$), but in children this difference was not significant ($\beta = 0.08$, $SE = 0.38$, $z = 0.22$, $p = .82$, $M_{Visual} = 0.23$, $M_{NoVisual} = 0.24$). The model did not yield any other significant interactions.

2.3. Discussion

Results from Experiment 1 show that both 4- to 5-year-old children and adults made generic adjustments by mentioning atypical/unusual instruments more often than typical instruments (even though this effect was more pronounced in adults). This extends prior findings by Brown and Dell (1987) and Lockridge and Brennan (2002) concerning typicality-based adjustments in story-retelling to spontaneous production with both adults and children.

Results also showed that adults made a particular-listener adjustment by mentioning instruments more frequently when the listener could not see the events but children's instrument mentions were unaffected by the listener's visual access. Interestingly, adults' adjustments to listener's visual access in our study was distinctly different from prior work. Contrary to Lockridge and Brennan, the increase in instrument mention when the listener lacked visual access to the events in our experiment was not affected by whether the instrument was atypical or not (i.e., by whether the listener would be able to infer it or not). It is possible that this finding represents a "coarse" audience design effect (as opposed to the more specific one found by Lockridge & Brennan), which arose as a result of using a confederate-listener instead of a naïve participant. Alternatively, this finding may be due to the nature of our task; asking participants to describe single instrument events in our study elicited more instrument mentions (both typical and atypical) than asking participants to narrate multi-event stories in both prior studies. The fact that atypical instrument mention was already very high when the listener could see the events may

Table 3
Parameter estimates for instrument mention in Experiments 1, 2, and 3 (see Appendix B for random effects)

Fixed Effects	Estimate	*SE*	*z*
Experiment 1			
Intercept	−0.38	0.36	−1.06
Typicality (Typical vs. Atypical)	3.51	0.69	5.12***
Age (Children vs. Adults)	−3.38	0.33	−10.19***
Access (Visual vs. No Visual)	0.79	0.30	2.65**
Typicality (Typical vs. Atypical): Age (Children vs. Adults)	−1.73	0.56	−3.07***
Typicality (Typical vs. Atypical): Access (Visual vs. No Visual)	0.28	0.47	0.60
Age (Children vs. Adults): Access (Visual vs. No Visual)	−1.42	0.59	−2.39*
Typicality (Typ vs. Atyp): Age (Ch vs. Ad): Access (V vs. NV)	−0.52	0.94	−0.55
Experiment 2			
Intercept	−0.29	0.32	−0.90
Typicality (Typical vs. Atypical)	3.12	0.60	5.24***
Age (Children vs. Adults)	−3.27	0.33	−9.89***
Access (Visual vs. No Visual)	0.86	0.29	2.93**
Typicality (Typical vs. Atypical): Age (Children vs. Adults)	−2.21	0.57	−3.89***
Typicality (Typical vs. Atypical): Access (Visual vs. No Visual)	−0.12	0.47	−0.25
Age (Children vs. Adults): Access (Visual vs. No Visual)	−1.54	0.59	−2.63**
Typicality (Typ vs. Atyp): Age (Ch vs. Ad): Access (V vs. NV)	−1.43	0.95	−1.54
Experiment 3			
Intercept	0.08	0.35	0.22
Typicality (Typical vs. Atypical)	2.79	0.57	4.86***
Age (Children vs. Adults)	−3.38	0.30	−11.36***
Access (Visual vs. No Visual)	0.31	0.26	1.22
Typicality (Typical vs. Atypical): Age (Children vs. Adults)	−1.60	0.52	−3.09**
Typicality (Typical vs. Atypical): Access (Visual vs. No Visual)	0.13	0.42	0.30
Age (Children vs. Adults): Access (Visual vs. No Visual)	−0.61	0.51	−1.19
Typicality (Typ vs. Atyp): Age (Ch vs. Ad): Access (V vs. NV)	−0.98	0.84	−1.18

Significance levels: $*p < .05$, $**p < .01$, $***p < .001$.

have prevented an asymmetric increase in the mention of atypical instruments when the listener could not see the events (see Table 2).[3] This possibility is supported by the observation that the numerical trend was the opposite than expected, such that there was an increase in the mention of typical rather than atypical instruments by adult speakers when the listener had no visual access to the events (Typical instruments: $M_{Visual} = 0.29$ vs. $M_{NoVisual} = 0.46$; Atypical instruments: $M_{Visual} = 0.88$ vs. $M_{NoVisual} = 0.97$).

Importantly, children's typicality-based adjustments did not seem to be driven by considerations of a particular addressee. One possible explanation for this result is that children had difficulty estimating the goals of the communicative exchange (e.g., see Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Papafragou & Musolino, 2003, for similar arguments); asking children to simply describe events for a silent listener may not have highlighted the listener's specific communicative needs. Experiment 2 explores this possibility by clarifying the listener's need for specific, detailed information about the events.

3. Experiment 2

Experiment 2 introduced a single change to the basic paradigm of Experiment 1: Participants were asked to describe the events so that an adult listener (a confederate of the experimenter's) could draw the events on a sketchpad. We reasoned that, given the addressee's goal, children would need to produce more complete event descriptions (including more instruments) in the No Visual compared to the Visual Access condition. (Adults were already sensitive to Visual Access in the more neutral Experiment 1.)

3.1. Method

3.1.1. Participants

Participants were 48 4- to 5-year-old children and 48 adults. The children were between the ages of 4;3 to 5;10, with a mean age of 4;10. All children were recruited from and tested at daycares in Newark, Delaware. Adults were undergraduate students at the University of Delaware and received course credit for their participation. None of these participants had taken part in Experiment 1.

3.1.2. Materials

Materials were the same as in Experiment 1.

3.1.3. Procedure

Procedure was as in Experiment 1 except that participants were asked to describe the events to an adult listener (a confederate of the experimenter) who had to draw the events on a sketchpad on the basis of how participants described them. The drawings of the "listener" were shown to participants only at the end so as not to affect their responses.

Table 4
Proportion of typical and atypical instrument mention for each age group and visual access condition in Experiment 2

	Children				Adults			
	Visual Access		No Visual Access		Visual Access		No Visual Access	
Category	Typ	Atyp	Typ	Atyp	Typ	Atyp	Typ	Atyp
Explicit mention								
Within Clause								
After the verb	0.03	0.23	0.03	0.14	0.15	0.70	0.34	0.73
Before the verb	0.01	0.02	0	0.01	0.03	0.06	0.03	0.11
Incorporated into the verb	0.06	0.01	0.07	0.03	0.06	0.03	0.04	0.02
Total	0.10	0.26	0.10	0.18	0.24	0.79	0.41	0.85
Separate Clause								
After the verb	0	0.03	0.01	0.01	0.01	0.03	0.02	0.09
Before the verb	0	0.01	0	0.03	0	0.01	0.01	0.03
Total	0	0.04	0.01	0.04	0.01	0.04	0.03	0.12
Implicit mention	0	0.06	0	0.06	0	0.01	0	0
Other	0	0.01	0.02	0.04	0	0	0	0
Total mention	0.10	0.37	0.13	0.32	0.25	0.84	0.45	0.97

3.2. Results

We analyzed Instrument Mention with a model that included Typicality (Typical, Atypical), Age (Children, Adults), Visual Access (Visual Access, No Visual Access), and all their interactions as fixed predictors. The model also included random by-Participants intercepts and slopes for Typicality, as well as random by-Item intercepts and slopes for Typicality. The data are summarized in Fig. 1 and Table 4. Table 3 presents the parameter estimates for the multi-level model of instrument mention (see Appendix B for random effects). Results largely replicated those of Experiment 1: The model yielded significant effects of Typicality ($M_{Atypical} = 0.63$, $M_{Typical} = 0.23$), Age ($M_{Adults} = 0.64$, $M_{Children} = 0.27$), and Visual Access ($M_{Visual} = 0.39$, $M_{NoVisual} = 0.47$), which were qualified by an interaction between Typicality and Age and an interaction between Age and Visual Access. Follow-up analyses showed that the Typicality by Age interaction was due to the fact that, although atypical instruments were used more frequently than typical instruments by both children and adults, the difference was greater in adults ($\beta = 4.72$, $SE = 0.85$, $z = 5.58$, $p < .001$, $M_{Atypical} = 0.92$, $M_{Typical} = 0.38$) than in children ($\beta = 2.07$, $SE = 0.77$, $z = 2.70$, $p = .007$, $M_{Atypical} = 0.37$, $M_{Typical} = 0.11$). The interaction between Age and Visual Access was due to the fact that adults mentioned instruments more frequently when the listener could not see the events compared to when she could see the events ($\beta = 1.83$, $SE = 0.60$, $z = 3.07$, $p = .002$, $M_{Visual} = 0.56$, $M_{NoVisual} = 0.72$), but there was no such difference in children ($\beta = 0.11$, $SE = 0.31$, $z = 0.35$, $p = .72$, $M_{Visual} = 0.23$, $M_{NoVisual} = 0.23$).

3.3. Discussion

Results of Experiment 2 replicated those of Experiment 1. Specifically, adults made both generic-listener adjustments (based on the typicality of instruments used in the stimuli) and particular-listener adjustments (based on whether the addressee had visual access to the events or not). Children also made generic, typicality-driven adjustments but failed to mention instruments more frequently when the addressee lacked visual access to the events. Thus, the presence of a stated goal for the addressee did not alter the basic pattern of results.

It is possible that despite the clear communicative advantages that knowing about instruments held for the addressee, her informational needs were not sufficiently transparent to children. Experiment 3 introduced modifications to Experiment 2 so as to make the task more interactive and the informational needs of the addressee explicit.

4. Experiment 3

Experiment 3 made three main modifications to the paradigm of Experiment 2. First, the addressee explicitly expressed the need for help in order to fulfil her goal (draw accurate pictures). Prior work with children has shown that "helping" contexts are more motivating for children and more likely to make them take into account other people's knowledge and beliefs (e.g., Buttelmann, Carpenter, & Tomasello, 2009). Second, the visual perspective of the listener was made explicit in order to further clarify the link between what the addressee could see and how much information she needed. Children have been shown to have important limitations in drawing that link (see Moll & Kadipasaoglu, 2013, for a review). Finally, the addressee explicitly asked for information before each trial in a way that highlighted her informational needs at each point in the discourse. We hypothesized that these modifications would encourage children to take into account the addressees' informational needs, since other studies using engaged listeners—even confederates of the experimenter's—have elicited messages tailored to the needs of the exchange (e.g., Bahtiyar & Küntay, 2009; Grigoroglou & Papafragou, 2019; Nadig & Sedivy, 2002; Nilsen & Graham, 2009; see also Kuhlen & Brennan, 2013, for a discussion). We also hypothesized that this paradigm would lead to increased instrument mentions (and probably greater overall informativeness) from both adults and children compared to our previous, less interactive experiments.

4.1. Method

4.1.1. Participants

Participants were 48 4- to 5-year-old children and 48 adults. The children were between the ages of 4;1 to 5;11, with a mean age of 4;9. All children were recruited from and tested at daycares in Newark, Delaware, and in a local children's museum. Adults were undergraduate students at the University of Delaware and received course credit for their participation. None of these participants had taken part in the previous experiments.

4.1.2. Materials

Materials were the same as in Experiments 1–2.

4.1.3. Procedure

The procedure was similar to Experiment 2 with the changes described below. Experiment 3 had an introductory and a main experimental phase. As in Experiments 1–2, the experimenter initially introduced the participant and the addressee to each other and assigned appropriate seats: Participants were seated in front of a computer screen and the addressee either sat next to the participant so that she could also see the screen (*Visual Access* condition) or across from the participant, so that she could not see the screen (*No Visual Access* condition). Unlike the previous experiments, the visual barrier used in the No Visual Access condition was not yet placed between the participant and the addressee in order to allow visual contact between them during the introductory phase of the experiment. As in Experiment 2, participants were asked to describe the videos to the addressee (an adult confederate) so that she could draw pictures on a sketchpad. Unlike Experiment 2, the addressee was instructed to act surprised when she was told she had to draw pictures and to confess that she was not very good at drawing. The experimenter explained that the addressee should not worry because the participant would help her draw "nice pictures." Children were then explicitly asked if they were willing to help the addressee (they all helpfully agreed; when testing adults, this step was omitted).

Next, the experimenter drew participants' attention to the addressee's visual access by asking both interlocutors if they could see the screen (*Visual Access* condition) or by placing an opaque barrier between the two interlocutors to "make sure that the game is fair and [the addressee] cannot see the computer screen" (*No Visual Access* condition). Addressees were instructed to act naturally during this phase and converse freely with the participant and the experimenter to ensure authenticity in the interaction between interlocutors. Such measures have been shown to overcome problems with the use of confederates in production studies (see Kuhlen & Brennan, 2013, for a discussion).

In the main experimental phase, unlike the previous two experiments, the addressee had a more active role: At the beginning of each trial, the addressee initiated communication by saying, "Tell me, what do you see?" Following the participant's response, the addressee replied, "Alright, let's see, how can I draw this . . .?", and started drawing the picture. At the end of the trial, the addressee expressed her readiness for the next trial by saying, "Ok, I hope this looks good! I am ready for the next one." As in Experiment 2, participants were not allowed to see the addressee's drawings until the end of the experiment.

4.2. Results

4.2.1. Mention of instruments

We analyzed Instrument Mention with a model that included Typicality (Typical, Atypical), Age (Children, Adults), Visual Access (Visual Access, No Visual Access), and all their interactions as fixed predictors. The model also included random by-Participants intercepts and slopes for Typicality, as well as random by-Item intercepts and slopes for Typicality. The data are summarized in Fig. 1 and Table 5. Table 3 presents the parameter estimates for the multi-level model of Instrument Mention (see Appendix B for random effects). The model showed a significant effect of Typicality, with atypical instruments being mentioned more frequently than typical instruments ($M_{Atypical} = 0.66$, $M_{Typical} = 0.32$). There was also a main effect of Age, with adults mentioning instruments more frequently than children ($M_{Adults} = 0.71$, $M_{Children} = 0.27$). These effects were qualified by an interaction between Typicality and Age. Follow-up analyses showed that this interaction was due to the fact that atypical instruments were mentioned more frequently than typical instruments by both age groups but the difference was greater in adults ($\beta = 3.14$, $SE = 0.55$, $z = 5.70$, $p < .001$, $M_{Atypical} = 0.93$, $M_{Typical} = 0.50$) than in children ($\beta = 2.46$, $SE = 0.89$, $z = 2.76$, $p = .006$, $M_{Atypical} = 0.39$, $M_{Typical} = 0.13$). Visual Access did not have a significant effect on instrument mention ($M_{Visual} = 0.47$, $M_{NoVisual} = 0.51$). No other effects or interactions were significant.

4.2.2. *Comparison across experiments*

To test whether the degree of the listener's involvement in the task affected the detail of participants' descriptions, we compared instrument mention across the three experiments. We hypothesized that participants would be more likely to add information about instruments for an interactive listener who is explicitly asking for information (Experiment 3) compared to a less actively engaged listener (Experiments 1–2). We also hypothesized that speakers would offer more instruments for a listener with a

Table 5
Proportion of Typical and Atypical Instrument Mention for each Age Group and Visual Access condition in Experiment 3

	Children				Adults			
	Visual Access		No Visual Access		Visual Access		No Visual Access	
Category	Typ	Atyp	Typ	Atyp	Typ	Atyp	Typ	Atyp
Explicit mention								
Within Clause								
After the verb	0.03	0.17	0.05	0.17	0.22	0.53	0.22	0.54
Before the verb	0	0.03	0	0.02	0.04	0.11	0.02	0.18
Incorporated into the verb	0.04	0.04	0.06	0.02	0.05	0.02	0.07	0.03
Total	0.07	0.24	0.11	0.21	0.31	0.66	0.31	0.75
Separate Clause								
After the verb	0	0.01	0.01	0.02	0.07	0.08	0.07	0.02
Before the verb	0.01	0.06	0.02	0.03	0.09	0.16	0.15	0.18
Total	0.01	0.07	0.3	0.05	0.16	0.24	0.22	0.20
Implicit mention	0	0.05	0	0.10	0	0.01	0	0
Other	0.04	0.05	0.01	0	0	0	0	0
Total mention	0.12	0.42	0.15	0.37	0.48	0.91	0.53	0.95

clear goal (Experiment 2) than a listener with a less specified goal (Experiment 1). To test these predictions, we used a model that included Typicality (Typical, Atypical) as a first-level predictor and Age (Children, Adults), Visual Access (Visual Access, No Visual Access), and Experiment (1, 2, 3) as second-level predictors as well as their interactions. The model also included random by-Participants intercepts and slopes for Typicality, as well as random by-Item intercepts and slopes for Typicality. The fixed effect of Experiment was analyzed with two simple contrasts comparing the more interactive Experiment 3 to the less interactive Experiments 1 and 2 combined (c_1: 0.33, 0.33, −0.66) and Experiment 1 to Experiment 2 (c_2: −0.5, 0.5, 0). Unsurprisingly given our previous analyses, the model showed significant effects of Typicality ($\beta = 3.12$, $SE = 0.39$, $z = 8.01$, $p < .001$), Age ($\beta = -3.29$, $SE = 0.18$, $z = -17.99$, $p < .001$) and Visual Access ($\beta = 0.65$, $SE = 0.16$, $z = 3.98$, $p < .001$), qualified by significant interactions between Typicality and Age ($\beta = -1.84$, $SE = 0.31$, $z = -5.93$, $p < .001$) and Age and Visual Access ($\beta = -1.19$, $SE = 0.33$, $z = -3.65$, $p < .001$). Importantly, the model also returned a significant effect of Experiment, with participants of Experiment 3 being more likely to mention instruments than participants of Experiments 1 and 2 combined ($\beta = -0.43$, $SE = 0.18$, $z = -2.38$, $p = .02$, $M_3 = 0.49$, $M_{1\&2} = 0.44$). Instrument mention was no different in Experiments 1 and 2 ($\beta = 0.12$, $SE = 0.22$, $z = 0.57$, $p = .57$, $M_1 = 0.45$, $M_2 = 0.43$). No other effects or interactions were significant.

4.3. Discussion

The results of Experiment 3 show that, when children and adults communicated with an interactive addressee, they frequently made generic (typicality-based) adjustments but did not adjust to the visual access of their addressee. Different explanations seem to underlie this finding for each age group (see Fig. 1). Children in Experiment 3, similarly to the previous two experiments, simply ignored the addressee's particular needs as determined by visual perspective. Adults, however, became much more detailed in their descriptions in this more interactive experiment, independently of whether the addressee could see the events or not (cf. also Grigoroglou & Papafragou, 2019), possibly because they considered that offering additional, often redundant information to an addressee who was explicitly requesting it helped her fulfil her task (i.e., make accurate drawings). Consistent with this finding, prior work shows that cooperative adult speakers often provide over-informative descriptions to facilitate the achievement of the listener's conversational goals (e.g., Arts, Maes, Noordman, & Jansen, 2011; Maes, Arts, & Noordman, 2004; see also Davies & Katsos, 2016, for discussion). Interestingly, in our experiments, this effect arose when the addressee's informational needs were very explicit and regularly stated by herself (as in Experiment 3) but less so when these needs were introduced once by the experimenter (as in Experiment 2).

The comparison of instrument mention across experiments showed that participants in Experiment 3 mentioned instruments more frequently compared to the less interactive Experiments 1 and 2. Still, children's mention of instruments remained low even in the more interactive Experiment 3, and the difference between Experiment 3 and Experiments 1–2 was small, indicating that the effect of the interactive addressee on children's production was rather limited.

5. General discussion

Adjusting one's speech to common ground shared with others requires building and maintaining a model of the listener (Clark & Marshall, 1981). This study investigated children's and adults' linguistic choices when describing events to address this process of audience design. Inspired by adult psycholinguistic research (beginning with Brown & Dell, 1987), we sought to test the hypothesis that distinct types of audience design (generic vs. particular, and subtypes thereof) should differentially affect children's linguistic choices during production.

This hypothesis was confirmed in our data. In Experiments 1–3, both 4- to 5-year-old children and adults made generic-listener adjustments by mentioning highly unpredictable (i.e., atypical) instruments more frequently than typical instruments. Furthermore, adults performed particular-listener adjustments by mentioning instruments more frequently when the listener could not see the events (as opposed to when the listener could see the events) but children's mentions of instruments were unaffected by the listener's visual access. Although children did not make particular adjustments to their listener's visual access, they did make some adjustments to their listener's communicative profile: as a comparison across experiments showed, children were more likely to add specific information about instruments when communicating with an interactive addressee (Experiment 3) compared to a silent addressee (Experiments 1–2). Even so, children's instrument mentions remained particularly low across all experiments, and their descriptions remained underinformative given the hearer's stated goal (i.e., to make accurate drawings).

5.1. Audience design in children and adults

Our findings confirm and extend prior work (e.g., Lockridge & Brennan, 2002) by demonstrating that adult speakers readily engaged in audience design at multiple levels when communicating with an interlocutor. They also show that 4- and 5-year-old children's respective performance was limited: Across three experiments, children mentioned event components that would be unexpected or noteworthy for any comprehender (i.e., atypical instruments) but, unlike adults, failed to offer more instrument information to listeners with limited knowledge. This pattern can be explained by a distinction between "particular" and "generic" listener models within a more specific and nuanced account of children's audience design compared to past work (cf. also Grigoroglou & Papafragou, 2019). On this account, maintaining and updating highly "particular" and dynamic representations of common ground (i.e., present, moment-by-moment representations of an exchange), especially when these involve constantly suppressing one's own visual perspective to adopt the perspective of the interlocutor, pose different demands compared to "generic," more stable representations (e.g., knowledge shared in the community/generally known facts, or memories of physical or linguistic co-presence).

Specifically, the robust presence of the instrument typicality effect found in 4- and 5-year-old children's speech in our experiments can be explained by the fact that this adjustment was computationally simple, as it did not require children to monitor the beliefs of their particular interlocutor or contrast their knowledge to the listener's but to simply consider what is conventional within the community, themselves included (e.g., that eating spaghetti with a serving spoon is unconventional). By contrast, the absence of a visual access effect in children can be explained by the fact that this adjustment was costly to implement, as it required children in the No Visual Access condition to constantly keep in mind that their listener could not see the events (hence she did not know what instrument was used), contrast this information with their own privileged knowledge, and repeat this process for every trial. Therefore, maintaining a "particular" model of the listener, as opposed to a more "generic" one seems to be associated with additional costs. This analysis is consistent with findings showing that children with better cognitive skills (e.g., executive functioning, working memory, mentalizing skills) are also better communicators (e.g., Nilsen & Graham, 2009; Nilsen et al., 2015; Resches & Pérez Pereira, 2007; Wardlow & Heyman, 2016).

One of the most interesting findings of our study was that both young and more experienced (adult) speakers became more likely to offer instrument information when the listener was interactive and engaged (Experiment 3) compared to silent and passive (Experiments 1 and 2). Positing a goal for a silent addressee did not by itself affect how much instrument information speakers encoded in either age group (Experiment 2 vs. 1). This pattern is consistent with the idea that adults are more likely to make production adjustments when their listeners have (or are thought to have) actual informational needs (Buz et al., 2016; Kuhlen & Brennan, 2010; Lockridge & Brennan, 2002). It also comports with the observation that children are more likely to be adequately informative in studies where the listener is an actual person with "real" informational needs (whether a confederate of the experimenter's—Bahtiyar & Küntay, 2009; Grigoroglou & Papafragou, 2019; Nadig & Sedivy, 2002; Nilsen & Graham, 2009; or a naïve participant—Köymen et al., 2016; O'Neill, 1996; Perner & Leekam, 1986) as opposed to an imaginary addressee or no specific addressee at all (e.g., Davies & Katsos, 2010; Girbau, 2001; cf. Davies & Kreysa, 2018; Rabagliati & Robertson, 2017).

Notice that the increase in instrument mention in the interactive Experiment 3 was not affected by whether the listener had visual access to the events. This pattern was particularly striking in children who still massively failed to include instrument information in Experiment 3 (adults almost always added instrument information, as discussed earlier). We propose that adaptations to the communicative profile of the listener differ from adaptations to the listener's visual access, even though they both seem tied to a particular listener in an exchange. Tailoring a message to the listener's profile did not require complex mentalizing or inhibition, but merely involved considering the speaker's global goal (already known to both speaker and listener) to collect enough information to produce accurate drawings. Considering this goal could have been made once for the specific listener and did not require constant updating. This analysis is consistent with claims about simple, "one-bit" partner models, whereby audience design is more likely to occur when listener-specific information is clear, simple, and easy to compute (see Brennan, Galati, & Kuhlen, 2010, for a discussion). Interactivity might have made attending to or fulfilling this goal more compelling because of the increased rewards of social interaction (including the psychological benefit of "helping" someone else; cf. Hamann, Warneken, & Tomasello, 2012). Therefore, in terms of audience design demands, the adjustment to the listener's profile would fall in between a particular adjustment to the listener's visual perspective and a generic adjustment to what is commonly known in a community.

Overall, the present data and analysis suggest that children's ability to adjust their speech to the common ground shared with an addressee should not be taken to reflect an "all or nothing" ability but should be conceptualized as a set of distinct cognitive abilities whose degree of difficulty spans a continuum (see Liebal, Carpenter, & Tomasello, 2013; Matthews et al., 2006; Brown-Schmidt & Heller, 2018, for related views). There is some indication that this perspective also extends to adults. In our adult data (as in Lockridge & Brennan, 2002), the effect of visual access was smaller than the effect of typicality, suggesting that adjusting one's speech to the particular visual perspective of the addressee was costlier than adjusting one's speech to the model of a generic listener, even for adults. Furthermore, as the adult findings of Experiment 3 showed, considerations of how much information a listener needs based on her visual perspective are independent from other listener-particular considerations (i.e., the listener's communicative profile). Together, our adult and developmental findings are consistent with the possibility that representations of different aspects of common ground in memory are independent from each other (see Brown-Schmidt & Duff, 2016; Brown-Schmidt & Heller, 2018, for related discussion; and section 4.4 below).

5.2. An explanation of prior developmental findings

The present nuanced perspective on children's production adaptations can explain discrepant findings in prior developmental work. Recall that a puzzle in referential communication is that sometimes even very young children (2- to 3-year-olds, e.g., Matthews et al., 2006; O'Neill, 1996; Perner & Leekam, 1986) make successful particular-listener production adjustments, while at other instances, children show non-adult-like performance until a fairly late age (8- to 12-year-olds, for example, Deutsch & Pechmann, 1982; Fukumura, 2016). The puzzle can be resolved once it becomes clear that different studies have actually tested distinct particular-listener adaptations. For instance, studies with very young children typically do not require children to track another person's visual perspective in real time but rather to remember their "prior engagement" with this person (i.e., whether the person was present or absent when relevant information was presented). This type of shared knowledge with a partner seems to be available earlier and be cognitively less demanding for children (e.g., see Perner & Roessler, 2012; Moll, Richter, Carpenter, & Tomasello, 2008, for relevant evidence with infants) than the ability to understand how others might perceive things visually (see Moll & Kadipasaoglu, 2013). This latter ability has been argued to develop between the ages of 4 and 5 (Moll, Meltzoff, Merzsch, & Tomasello, 2013).

Even beyond the preschool years, visual perspective-taking abilities in communication may not always be readily implemented. Studies that show successful production adjustments with older children (5- to 6-year-olds, e.g., Bahtiyar & Küntay, 2009; Nadig & Sedivy, 2002; Nilsen & Graham, 2009) used specific techniques to reduce the cognitive load involved in keeping track of another person's perspective while suppressing one's own and holding that information in memory on a trial by trial basis. For instance, in Nadig and Sedivy (2002), before each trial, the confederate-listener was asked to close his eyes so that the experimenter could place the "secret" object in the compartment that was hidden from the listener's view (but visible to the child). After each trial, the child and the listener played "a guessing game to reinforce the fact that one of the objects was hidden from the confederate's view" (Nadig & Sedivy, 2002, p. 331). In our own studies, the No Visual Access condition did not include such constant reminders but required speakers to navigate the differences in the ways themselves and their interlocutor viewed visual stimuli in each individual trial.

5.3. Extensions of current study

The present paradigm could be extended in several directions to more fully replicate the dynamic of naturalistic conversations. First, even in the most interactive Experiment 3, the conversational goal between speaker and listener was not truly shared (i.e., making accurate drawings was exclusively the listener's goal). Future versions of this experiment could explore whether engagement in a truly joint activity could increase partner-particular adaptations, especially in children (see also Tomasello, 2007; Tomasello, Carpenter, Call, Behne, & Moll, 2005).

Second, and relatedly, our design did not allow for repairs or reformulations of under-informative utterances. In everyday speech production, constituents are monitored either before or after articulation (Levelt, 1989) and can be repaired if found inappropriate. Even though 4- to 5-year-olds do not typically engage in self-monitoring for ambiguity (and rarely act on it even if they do; Rabagliati & Robertson, 2017), when asked to give additional information or are given other evidence that their referential attempts have failed, children from at least the age of 2 can repair their utterances to facilitate the listener's comprehension (Deutsch & Pechmann, 1982; Matthews, Butcher, Lieven, & Tomasello, 2012; Matthews, Lieven, & Tomasello, 2007; Sarılar, Matthews & Küntay, 2013; Uzundağ & Küntay, 2018; see also Golinkoff, 1986; Nilsen & Mangal, 2012). It is an open possibility that children in our paradigm would have been more likely to make particular-listener adaptations if their interlocutor had provided more active evidence about what they understood.

Finally, our study differed from prior work that measured adjustments of the specificity of a referring expression in an already existing (obligatory) component (e.g., "the man" vs. "he," "the big glass" vs. "glass") since it required monitoring both instrument typicality and a partner's perspective to selectively add an (optional) syntactic constituent (e.g., a prepositional phrase) to one's sentence when appropriate. Furthermore, unlike the informativeness of referring expressions that is directly linked to considerations of common ground (Ariel, 1990), understanding that an unusual instrument cannot be inferred by a listener, especially one who lacks visual access to the scene, and should, therefore, be mentioned is a more complex skill, requiring the coordination of different sources of information. The fact that studies demonstrating early sensitivity to a partner's perspective in children's speech have typically focused on referring expressions supports this analysis (e.g., see Matthews et al., 2006). This line of reasoning raises the possibility that children might be more likely to satisfy the listener's informational needs if production costs were reduced. For instance, it would be interesting to test whether children would add information about atypical instruments to expand on someone else's underinformative description for the sake of a listener who lacks visual access to the event.

5.4. Final thoughts

Our results have both theoretical and methodological implications for future research on children's audience design. From a theoretical perspective, they suggest that the question "Do children adapt their speech to the needs of their addressees?" should be replaced by a set of more specific questions such as "What factors determine children's ability to adjust to their interlocutors' needs?" that can be further sub-divided to target particular- vs. generic-addressee-oriented adaptations (and their sub-types). The field could also benefit from further contact with adult psycholinguistics, where it has long been recognized that the process of designing messages with a particular addressee's needs in mind depends on several cognitive factors (see Konopka & Brown-Schmidt, 2014). For example, when under time pressure, adult speakers in referential communication tasks are less sensitive to the addressee's knowledge state (Horton & Keysar, 1996). Similarly, adults are less successful with audience design when they have to track perspectives of other participants in a conversation that are less distinct from each other (Heller, Skovbroten, & Tanenhaus, 2009; Horton & Gerrig, 2005) or when they have to suppress highly salient privately held information to adopt the addressee's perspective (Wardlow Lane & Ferreira, 2008; Wardlow Lane, Groisman, & Ferreira, 2006). A synthesis of adult and developmental findings would be useful here.

Methodologically, the present discussion suggests that the current lack of consensus about children's audience design might be due to the fact that prior experimental paradigms, although mostly focused on referential disambiguation, differed greatly in terms of their specific properties. Looking at how children freely describe events provides a novel area for studying children's communicative behavior, and it can be used to systematically explore sensitivity to different types of common ground assumptions (see also Köymen et al., 2016; Mammen, Köymen, & Tomasello, 2017).

Acknowledgments

We thank the children and the parents in the preschools of Newark, Delaware, and the Delaware Children's Museum (DCM). We thank Jessica Granieri, Megan Kelly, Sally McHugh, and Queen Ralf, for their assistance in these studies. Portions of these data have been presented at TiXPrag1, CogSci 2016, BUCLD 42, and IWLP 2018. This research has been partially supported by NSF grant BCS-0749870 to Anna Papafragou and a Doctoral Fellowship to Myrto Grigoroglou, while she was a graduate student at the University of Delaware.

Notes

1. In Brown and Dell's original scheme, implicit mention was used for cases where instruments were mentioned after the clause describing the main event, often at the very end of the story. This was because their task involved retelling stories, where the event with the typical/atypical instrument was only a part of the story. Our task involved the description of a single event, so this category was no longer relevant.
2. We coded such utterances as (implicit) instrument mention because their use, although limited, was systematic for specific atypical instrument events across different participants. Although we recognize that these locutions may not always succeed in specifying the exact identity of the atypical instrument that was used (e.g., "folding" may not necessarily invoke the use of a towel), they still make important semantic components of the instrument inferable (e.g., "folding" implies the use of a fabric-like entity). To ensure that this coding choice did not affect the findings, we analyzed our data without the implicit instrument mention category and the patterns of results did not differ.
3. To directly test whether our data provide evidence in favor of the null hypothesis (i.e., that there is no interaction of Typicality and Visual Access), we computed BIC-based Bayes factors (Wagenmakers, 2007). This analysis showed that the estimated Bayes factor in favor of H1 (i.e., the full model including the two-way interaction of Typicality and Visual Access and the three-way interaction of Typicality, Visual Access, and Age) over H0 (i.e., a model without these interactions) was 0.001, suggesting strong evidence for H0 (Jeffreys, 1961; Raftery, 1995). We also performed two additional comparisons of the full model (including all interactions of Typicality and Visual Access) and a model without the two-way interaction of Typicality and Visual Access and a model without the three-way interaction of Typicality, Visual Access, and Age. The comparison of the full model with the model without the two-way interaction yielded a Bayes factor of 1, suggesting no evidence in favor of H1. The comparison of the full model with the model without the three-way interaction yielded a Bayes factor of 0.034, suggesting strong evidence in favor of H0.

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. https://doi.org/10.1080/136828200247278

Allen, S. E. M., Hughes, M., & Skarabela, B. (2015). The role of cognitive accessibility in children's referential choice. In L. Serratrice & S. E. M. Allen (Eds.), *The acquisition of reference* (pp. 123–153). Amsterdam, The Netherlands: John Benjamins. https://doi.org/10.1075/tilar.15.06all

Allen, S., Skarabela, B., & Hughes, M. (2008). Using corpora to examine discourse effects in syntax. In H. Behrens (Ed.), *Trends in language acquisition research: Finding structure in data* (Vol. *6*, pp. 99–137). Amsterdam, The Netherlands: John Benjamins. https://doi.org/10.1075/tilar.6.07all

Andrés-Roqueta, C., & Katsos, N. (2017). The contribution of grammar, vocabulary and Theory of Mind in pragmatic language competence in children with autistic spectrum disorders. *Frontiers in Psychology*, *8*, 996. https://doi.org/10.3389/fpsyg.2017.00996

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London, UK: Routledge.

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, *23*(4), 495–527. https://doi.org/10.1080/01690960801920099.

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification facilitates object identification. *Journal of Pragmatics*, *43*(1), 361–374.

Baayen, R. H. (2008). *Analyzing linguistic data a practical introduction to statistics using R*. New York, NY: Cambridge University Press.

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. https://doi.org/10.1016/j.jml.2007.12.005

Bahtiyar, S., & Küntay, A. C. (2009). Integration of communicative partner's visual perspective in patterns of referential requests. *Journal of Child Language*, *36*(3), 529–555. https://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*(7), 1–13. https://doi.org/10.1177/0956797617699848

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. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://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*, *22*(6), 1482–1493. https://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. https://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*(4), 441–472. https://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*(5), 893–900. https://doi.org/10.3758/PBR.16.5.893

Brown-Schmidt, S., & Duff, M. C. (2016). Memory and common ground processes in language use. *Topics in Cognitive Science*, *8*(4), 722–736. https://doi.org/10.1111/tops.12224

Brown-Schmidt, S., & Heller, D. (2018). Perspective taking during conversation. In S. -A. Rueschemeyer & M. G. Gaskell(Eds), *Oxford handbook of psycholinguistics* (2nd ed.). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.013.23

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, *122*(2), 135–149. https://doi.org/10.1016/j.cognition.2011.10.002

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–342. https://doi.org/10.1016/j.cognition.2009.05.006

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language*, *89*(1), 68–86. https://doi.org/10.1016/j.jml.2015.12.009

Clark, E. V. (2007). Conventionality and contrast in language and language acquisition. *New Directions for Child and Adolescent Development*, *115*, 11–23.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. H. Joshi, B. I. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, UK: Cambridge University Press.

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua*, *120*(8), 1956–1972. https://doi.org/10.1016/j.lingua.2010.02.005

Davies, C., & Katsos, N. (2016). Reference and Informativeness as cognitive processes in verbal communication. In A. Rocci, & L. de Saussure (Eds.), *Verbal Communication. Handbooks of Communication Science* (pp. 123–140). Berlin, Germany: De Gruyter Mouton.

Davies, C., & Kreysa, H. (2018). Look before you speak: children's integration of visual information into informative referring expressions. *Journal of Child Language*, *45*(5), 1116–1143. https://doi.org/10.1017/S0305000918000120

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener". In D. Napoli, & J. Kegl (Eds.), *Bridges between psychology and linguistics* (pp. 105–129). San Diego, CA: Academic Press.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*(2), 159–184. https://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*(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002

Fukumura, K. (2016). Development of audience design in children with and without ASD. *Developmental Psychology*, *52*(1), 71–87. https://doi.org/10.1037/dev0000064

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51. https://doi.org/10.1016/j.jml.2009.09.002

Girbau, D. (2001). Children's referential communication failure: The ambiguity and abbreviation of message. *Journal of Language and Social Psychology*, *20*(1–2), 81–89. https://doi.org/10.1177/0261927X01020001004

Golinkoff, R. M. (1986). "I beg your pardon?": The preverbal negotiation of failed messages. *Journal of Child Language*, *13*, 455–76.

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. https://doi.org/10.1037/a0029467

Greenfield, P. M. (1980). Going beyond information theory to explain early word choice: A reply to Roy Pea. *Journal of Child Language*, *7*, 217–221. https://doi.org/10.1017/S0305000900007145

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

Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase children's informativeness in referential communication. *Developmental Psychology*, *55*(5), 951–966. https://doi.org/10.1037/dev0000693

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development*, *83*(1), 137–145. https://doi.org/10.1111/j.1467-8624.2011.01695.x

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: Shared knowledge affects referential form. *Topics in Cognitive Science*, *4*, 290–305. https://doi.org/10.1111/j.1756-8765.2012.01182.x

Heller, D., Skovbroten, K., & Tanenhaus, M. K. (2009). Experimental evidence for speakers' sensitivity to common vs. privileged ground in the production of names. Paper presented at the PRE-CogSci Workshop on the Production of Referring Expressions, Amsterdam, The Netherlands.

Hendriks, P. (2016). Cognitive modeling of individual variation in reference production and comprehension. *Frontiers in Psychology*, *7*, 506. https://doi.org/10.3389/fpsyg.2016.00506

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, *47*, 589–606.

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition*, *96*, 127–142. https://doi.org/10.1016/j.cognition.2004.07.001

Horton, W. S., & Keysar, B. (1996). When do speakers take into common ground? *Cognition*, *56*(1), 91–117.

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*(6), 597–626. https://doi.org/10.1023/a:1026468209238

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

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use. *Current Directions in Psychological Science*, *7*(2), 46–49. https://doi.org/10.1111/1467-8721.ep13175613

Konopka, A. E., & Brown-Schmidt, S. (2014). Message encoding. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 3–20). Oxford, UK: Oxford University Press.

Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology*, *52*(3), 423–429. https://doi.org/10.1037/dev0000089

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, *47*(7), 567–587. https://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*(1), 54–72. https://doi.org/10.3758/s13423-012-0341-8

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

Levinson, S. C. (1995). Three levels of meaning: Essays in honor of Sir John Lyons. In F. R. Palmer (Ed.), *Grammar and meaning* (pp. 90–115). Cambridge, UK: Cambridge University Press.

Liebal, K., Carpenter, M., & Tomasello, M. (2013). Young children's understanding of cultural common ground. *British Journal of Developmental Psychology*, *31*(1), 88–96. https://doi.org/10.1111/j.2044-835X.2012.02080.x.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, *9*(3), 550–557. https://doi.org/10.3758/BF03196312

Maes, A., Arts, A., & Noordman, L. (2004). Reference Management in Instructive Discourse. *Discourse Processes*, *37*(2), 117–144. https://doi.org/10.1207/s15326950dp3702_3.

Mammen, M., Köymen, B., & Tomasello, M. (2017). The reasons young children give to peers when explaining their judgments of moral and conventional rules. *Developmental Psychology*. Advance online publication. https://doi.org/10.1037/dev0000424

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*(2), 184–210. https://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*(3), 403–422. https://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*(6), 1744–1759. https://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. https://doi.org/10.1017/S0305000909009507

Moll, H., & Kadipasaoglu, D. (2013). The primacy of social over visual perspective-taking. *Frontiers in Human Neuroscience*, *7*(September), 1–9. https://doi.org/10.3389/fnhum.2013.00558

Moll, H., Meltzoff, A. N., Merzsch, K., & Tomasello, M. (2013). Taking versus confronting visual perspectives in preschool children. *Developmental Psychology*, *49*(4), 646–654. https://doi.org/10.1037/a0028633

Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what 'we' have shared in a special way. *Infancy*, *13*, 90–101. https://doi.org/10.1080/15250000701779402.

Moll, H., & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, *43*(2), 309–317. https://doi.org/10.1037/0012-1649.43.2.309

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329–336. https://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*(1), 55–78. https://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*(2), 220–249. https://doi.org/10.1016/j.cogpsych.2008.07.002

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*(5), 1121–1134. https://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. https://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*(2), 659. https://doi.org/10.2307/1131839

O'Neill, D. K. (2012). Components of pragmatic ability and children's language development. In H. -J. Schmid (Ed.), *Cognitive pragmatics* (pp. 261–288). Berlin: Mouton de Gruyter.

Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: evidence from event encoding. *Language, Cognition and Neuroscience*, https://doi.org/10.1080/23273798.2019.1589540 [advanced online publication].

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, *98*(3), B75–B87. https://doi.org/10.1016/j.cognition.2005.05.005

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, *86*(3), 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, *27*(1), 89–110. https://doi.org/10.1515/ling.1989.27.1.89.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Science*, *16*, 519–525. https://doi.org/10.1016/j.tics.2012.08.004.

Perner, J., & Leekam, S. R. (1986). Belief and quantity: Three-year olds' adaptation to listener's knowledge. *Journal of Child Language*, *13*(2), 305–315. https://doi.org/10.1017/S0305000900008072.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(02), 169–190; discussion 190–226. https://doi.org/10.1017/S0140525X04000056

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org

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. https://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.

Recanati, F. (2004). *Literal meaning*. Cambridge, UK: Cambridge University Press.

Resches, M., & Pérez Pereira, M. (2007). Referential communication abilities and Theory of Mind development in preschool children. *Journal of Child Language*, *34*, 21–52. https://doi.org/10.1017/S0305000906007641

Roberts, R. J., & Patterson, C. J. (1983). Perspective-taking and referential communication: The question of correspondence reconsidered. *Child Development*, *54*(4), 1005–1014. https://doi.org/10.2307/1129904

Salomo, D., Lieven, E., & Tomasello, M. (2013). Children's ability to answer different types of questions. *Journal of Child Language*, *40*(02), 469–491. https://doi.org/10.1017/S0305000912000050

Sarılar, A., Matthews, D., & Küntay, A. C. (2013). Hearing relative clauses boosts relative clause usage (and referential clarity) in young Turkish language learners. *Applied Psycholinguistics*, *36*(2), 175–202. https://doi.org/10.1017/S0142716413000192

Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication skills: The interaction of role-switching and difference rule training. *Journal of Experimental Child Psychology*, *38*(2), 191–207. https://doi.org/10.1016/0022-0965(84)90121-8

Sperber, D. (1994). Understanding verbal understanding. In J. Khalfa (Ed.), *What is intelligence?* (pp. 179–198). Cambridge, UK: Cambridge University Press.

Tomasello, M. (2007). Cooperation and communication in the 2nd year of life. *Child Development Perspectives*, *1*(1), 8–12. https://doi.org/10.1111/j.1750-8606.2007.00003.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(05), 675–735. https://doi.org/10.1017/S0140525X05000129

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*, 2–7. https://doi.org/10.1016/j.specom.2006.10.003

Uzundağ, A. B., & 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. https://doi.org/10.1016/j.jecp.2018.02.009

Vanlangendonck, F., Willems, R. M., Menenti, L., & Hagoort, P. (2016). An early influence of common ground during speech planning. *Language, Cognition and Neuroscience*, *31*(6), 741–750. https://doi.org/10.1080/23273798.2016.1148747

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

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. https://doi.org/10.1016/j.jecp.2016.05.016

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1466–1481. https://doi.org/10.1037/a0013353

Wardlow Lane, L., Groisman, M., & Ferreira, V. S. (2006). Don't talk about pink elephants! Speakers' control over leaking private information during language production. *Psychological Science*, *17*, 273–277. https://doi.org/10.1111/j.1467-9280.2006.01697.x

Whitehurst, G. J. (1976). The development of communication: Changes with age and modeling. *Child Development*, *47*(2), 473. https://doi.org/10.2307/1128804.

Whitehurst, G. J., & Sonnenschein, S. (1981). The development of informative messages in referential communication: Knowing when vs. knowing how. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 127–141). New York: Academic Press.

Wu, S., & Keysar, B. (2007). The effect of information overlap on communication effectiveness. *Cognitive Science*, *31*, 169–181. https://doi.org/10.1080/03640210709336989

Yoon, S. O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals on reference production in conversation. *Psychonomic Bulletin & Review*, *19*(4), 699–707. https://doi.org/10.3758/s13423-012-0262-6

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

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

Appendix A: Test stimuli

	Event	Atypical Instrument	Typical Instrument
1.	A man is painting a window	Using a paintbrush	Using a toothbrush
2.	A man is fixing a fence	Using a hammer	Using a shoe (as a hammer)
3.	A man is blowing his nose	Using tissue	Using a shirt
4.	A man is catching a ball	Using his hands	Using a basket
5.	A man is wrapping a present	Using paper	Using a towel
6.	A man is watering plants	Using a watering can	Using a hat
7.	A man is eating spaghetti	Using a fork	Using a large spoon
8.	A man is opening the door	Using a key	Using a hanger
9.	A man is opening a can	Using a can opener	Using a knife
10.	A man is digging a hole	Using a shovel	Using a plate (as a shovel)
11.	A man is cleaning the counter	Using a cloth	Using a stuffed animal
12.	A man is brushing his teeth	using a toothbrush	Using his finger

Appendix B: Random effects of Instrument Mention for the logistic mixed effects analyses in Experiments 1, 2, and 3

Random Effects	Variance
Experiment 1	
Participants (Intercept)	0.91
Typicality	0.52
Items (Intercept)	1.03
Typicality	4.01
Experiment 2	
Participants (Intercept)	0.91
Typicality	0.63
Items (Intercept)	0.90
Typicality	3.38
Experiment 3	
Participants (Intercept)	0.63
Typicality	0.35
Item (Intercept)	1.21
Typicality	3.23

Note: Maximal random effects structure justified by the design was used in all models. When the maximal model did not converge, the random component with the least variance was removed and the model was refit.