Journal of Experimental Psychology: Learning, Memory, and Cognition

Referential Processing in 3- and 5-Year-Old Children Is Egocentrically Anchored

Ekaterina Ostashchenko, Gaétane Deliens, Philippine Geelhand, Julie Bertels, and Mikhail Kissine

Online First Publication, October 4, 2018. http://dx.doi.org/10.1037/xlm0000659

CITATION

Ostashchenko, E., Deliens, G., Geelhand, P., Bertels, J., & Kissine, M. (2018, October 4). Referential Processing in 3- and 5-Year-Old Children Is Egocentrically Anchored. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. http://dx.doi.org/10.1037/xlm0000659

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/18/$12.00

http://dx.doi.org/10.1037/xlm0000659

Referential Processing in 3- and 5-Year-Old Children Is Egocentrically Anchored

Ekaterina Ostashchenko, Gaétane Deliens, Philippine Geelhand, Julie Bertels, and Mikhail Kissine
Université libre de Bruxelles

An ongoing debate in the literature on language acquisition is whether preschool children process reference in an egocentric way or whether they spontaneously and by-default take their partner's perspective into account. The reported study implements a computerized referential task with a controlled trial presentation and simple verbal instructions. Contrary to the predictions of the partner-specific view, entrained referential precedents give rise to faster processing for 3- and 5-year-old children, independently of whether the conversational partner is the same as in the lexical entrainment phase or not. Additionally, both age groups display a processing preference for the interaction with the same partner, be it for new or previously used referential descriptions. These results suggest that preschool children may adapt to their conversational partner; however, partner-specificity is encoded as low-level auditory-phonological priming rather than through inferences about a partner's perspective.

Keywords: language processing, lexical entrainment, referential communication, pragmatics, perspective taking

Virtually any object or individual can be referred to by more than one lexical description. Each of such referential alternatives corresponds to a conceptual perspective, as the selection of this or that description from a larger pool is driven by the interactional context and communicative goals. A crucial factor in such referential choices is the alignment of conversational partners on a way they designate a given referent. For instance, if in a conversation one speaker uses *silver band* to designate an item that could be referred to also as *metallic circle*, her interlocutor is likely to reuse the same description later in the conversation. In other words, conversational participants implicitly agree on a specific label for a recurring reference and expect this label to be reused consistently throughout the interaction. The recurring lexical label is thus said to be *entrained* during the conversation (Garrod & Anderson, 1987). Although lexical entrainment clearly contributes to conversational efficiency, a hotly debated, still unresolved issue is whether this process is grounded in adopting one's partner perspective or whether, instead, it instantiates an egocentric conversational strategy.

According to the first, "partner-specific" view, referential choices largely depend on one's conversational partner and the exchange dynamics (Brennan & Clark, 1996; Brown-Schmidt, 2009a; Metzing & Brennan, 2003). For the proponents of these models, conversational partners continuously monitor each other's cognitive perspectives and adjust their utterances accordingly. Mutually shared informational background would thus be determinant for lexical choices and for their entrainment. In these theoretical paradigms, language processing, as well as production, would be inherently partner-specific. Conversational partners would expect—in a way consistent with Grice's (1975) principle of cooperativity (and more precisely with his manner maxim)—each other to follow precedents, because it would be highly uncooperative for a speaker to shift away, for no reason, from a previously entrained way to refer to a given object.

The opposite, "egocentric" view is that lexical entrainment is partner-independent; rather than being determined on common ground, it would rely on the egocentric availability of the precedent (Barr & Keysar, 2002; Epley, Morewedge, & Keysar, 2004; Keysar, Barr, Balin, & Brauner, 2000; Kronmüller & Barr, 2006). Authors defending the egocentric view do not necessarily deny the existence of partner-specific effects but contend that one's partner perspective is only considered during the late stages of processing. Conversations would by default start off from an egocentric perspective—even though this initial egocentric viewpoint may be overridden at a later stage, in case adapting to the partner's perspective is required for successful communication (Barr &

Ekaterina Ostashchenko, Gaétane Deliens, and Philippine Geelhand, ACTE—Autism in Context: Theory and Experiment, LADISCO—Centre of Research in Linguistics, Université libre de Bruxelles Neurosciences Institute, Université libre de Bruxelles; Julie Bertels, CRCN - Center for Research in Cognition and Neurosciences, Université libre de Bruxelles Neurosciences Institute, LCFC - Laboratoire de Cartographie Fonctionnelle du Cerveau, Hôpital Erasme, Université Libre de Bruxelles; Mikhail Kissine, ACTE—Autism in Context: Theory and Experiment, LADISCO—Centre of Research in Linguistics, Université libre de Bruxelles Neurosciences Institute, Université libre de Bruxelles.

Correspondence concerning this article should be addressed to Ekaterina Ostashchenko, ACTE—Autism in Context: Theory and Experiment, Université Libre de Bruxelles, Avenue F.D. Roosevelt 50, B-1050 Brussels, Belgium. E-mail: Ekaterina.Ostashchenko@ulb.ac.be

Keysar, 2002; Deliens, Antoniou, Clin, & Kissine, 2017; Epley et al., 2004; Kronmüller & Barr, 2006; Kronmüller, Noveck, Rivera, Jaume-Guazzini, & Barr, 2017). In sum, for the proponents of this second position, conversational efficiency of keeping lexical choices constant is primarily egocentric and independent of the conversational partner's perspective or, for that matter, his or her identity.[1]

The debates between these models revolve around ample evidence that adults can easily track their partner's perspective, as this tracking is very flexible and not cognitively taxing. In many situations, common ground is automatically taken into consideration by partners, even though, in some circumstances, such as high cognitive load or low motivation, adults do still apply an exclusively egocentric strategy (Brown-Schmidt, 2009b; Cane, Ferguson, & Apperly, 2017).

Developmental literature offers less clear evidence on whether children systematically use perspectival information when interpreting language. Numerous studies on production of referential descriptions in preschoolers have shown that young children frequently fail to provide enough information for their partner to identify the intended referent. For example, in a study by Pechmann and Deutsch (1982) almost half of 7-year-old children failed to provide referential descriptions that were sufficiently unambiguous from their partner's perspective. For instance, they described an object by its color (e.g., "the red one"), even when several other objects of the same color were present. In a simplified version of the referential communication task—with only one descriptive dimension (e.g., size), instead of two (e.g., size and color)—children do display sensitivity to the common ground and produce disambiguating adjectives (Pechmann & Deutsch, 1982). However, their disambiguation strategy remains significantly less efficient than that in adults (Epley et al., 2004; Fukumura, 2016). Likewise, even though by the age of 6 children can take shared background into consideration during simple tasks, their performance rapidly reverts to egocentric once there is variation in stimuli array. Moreover, children produce uninformative or redundant messages as readily as informative ones, showing lack of adjustment to their partner (Whitehurst & Sonnenschein, 1978).

That said, Rabagliati and Robertson (2017) recently found that 3- to 5-year-old children do track their own descriptions for ambiguity; however, they do so after having produced these descriptions. This a posteriori self-monitoring probably evolves into a proactive self-monitoring of ambiguity avoidance at a preutterance stage. In other words, whereas adults automatically tune their production to the interlocutor before speaking, children may need additional training to produce descriptions adapted to the partner's informational needs. Consistently, even very young children (2 to 4 years old) become more efficient at producing informative descriptions when provided feedback about their uninformative attempts (Matthews, Lieven, & Tomasello, 2007). A number of other converging findings also suggest precocious sensitivity to the partner's perspective in preschoolers (Low, Drummond, Walmsley, & Wang, 2014; Onishi & Baillargeon, 2005; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010; Schneider, Bayliss, Becker, & Dux, 2012; Southgate, Senju, & Csibra, 2007).

It thus appears, on the one hand, that preschool children may track information related to their partner's perspective in many cases, but, on the other hand, they often fail to use it in referential tasks to produce informative descriptions. One plausible explanation is that tracking one partner's perspective is not directly inherent in referential and, more broadly, language processing. If perspective-taking is cognitively (and thus developmentally) distinct from language skills, it would make sense that children exhibit an egocentric strategy in referential communication tasks, failing to systematically heed the common ground.

There is no firm evidence on whether children automatically consider their conversational partner's perspective when they process referential information similarly to adults or whether they tend to interpret linguistic information in a more egocentric fashion. Adult-like processing of lexical entrainment by children would provide strong evidence in favor of a partner-specific account, indicating that lexical choices in communication are determined by the common ground. On a stronger interpretation, young children's early sensitivity to their partner's perspective would suggest that expectations about partner's intent is part and parcel of language production and comprehension (Tomasello, 1999). By contrast, if lexical entrainment in young children turns out to be egocentrically anchored, perspective-taking would likely be an optional pragmatic process, whose developmental course is at least in part language-independent.

As mentioned in the preceding text, lexical entrainment tasks offer a particularly promising window on the role of common ground in language processing, which may help clarify the developmental picture. Once a particular description (e.g., *silver band*) is associated with a corresponding referent, a precedent is established. If later this description is replaced by another one (e.g., *metallic circle*), the precedent is broken. To repeat, according to the partner-specific account, listeners will be less surprised when such modification is introduced by a new speaker than by the speaker who originally established a precedent. According to a partner-independent account, by contrast, the effect of precedent should be independent from the speaker's identity.

In studies with adults, researchers have found that common ground information guides comprehension of precedents (Brown-Schmidt, 2009a; Metzing & Brennan, 2003). Importantly, this conclusion rests on the presence of speaker-related effects both for maintained and broken precedents. Not only were the adult participants faster in identifying referents for maintained precedents with the same speaker, they were also more delayed when a precedent was broken by the speaker who introduced it versus when it was broken by a new speaker who was not privy to the lexical item that was previously entrained.

To the best of our knowledge, there are only two studies that explored similar partner-specific effects on lexical entrainment in young children. In Graham, Sedivy, and Khu (2014), 4-year-old children had to point to one of two objects described with an adjective followed by a noun (e.g., *the striped ball*). Behavioral observations of pointing to items were coupled with eye-tracking recordings. During the entrainment phase, experimenters established a referential precedent for each object; in the test phase, this precedent was either violated (e.g., *the yellow ball* instead of *the*

[1] A third approach by Heller, Parisien, and Stevenson (2016) suggests that both the egocentric perspective and common ground are integrated within referential processing. According to this Bayesian model of reference comprehension, listeners simultaneously consider evidence from both the egocentric perspective and common ground when developing expectations for the intended referent.

striped ball during warm-up) or maintained by the same or a new experimenter. Accuracy in children's pointing was significantly lower when the precedent was broken, irrespective of the partner's identity. Eye-movements data showed an advantage in early fixations toward the target item in the same partner condition. That is, children were faster in detecting the referent in the same partner condition, when they heard a description that has already been used by this partner (although the effect was short-lived and occurred only during the noun region). However, such facilitation effects in finding an item when precedents are maintained by the original speaker can be explained by memory-based mechanisms such as episodic priming and encoding specificity (Kronmüller & Barr, 2015). These mechanisms presumably are rooted in egocentric heuristics and do not require taking into consideration the conversational partner's perspective. Graham et al. (2014) did not report any statistically significant differences in eye-movement relative to a broken precedent by a new partner versus the same partner, which would have been indicative of a partner-specific processing. In the absence of a partner-specific processing of broken precedents, faster processing of maintained precedents with the same partner does not allow to conclude that lexical processing is guided by expectations about the partner's perspective in 4 year olds.

Matthews, Lieven, and Tomasello (2010) adopted the paradigm used with adults by Metzing and Brennan (2003). Children were asked to locate targets on a Plexiglas shelf following the instructions provided by an adult. After an entrainment phase, during which an adult established a referential precedent (e.g., *tree*), this precedent was broken (e.g., the word *bush* was used instead of *tree*) either by the same adult or a new one. Matthews and colleagues report that children's RTs were considerably slower in the same partner condition. Because the new adult is not supposed to have knowledge of the previously established referential precedent, this result indicates, according to the authors, sensitivity to the common ground. Recall that on an egocentric model one should expect violations of referential precedents to cause response delays irrespective of the speaker's status (same vs. new). Given the young age of their participants, the Matthews et al. (2010) study seems to provide decisive support for the partner-specific view or, at least, for the idea that perspective-shifting meshes with linguistic processing from very early acquisition stages. This ingenious paradigm has the further advantage to rely on referential descriptions that are already part of the children's lexicon, thus maximally isolating perspective-shifting from other language related processing factors.

That said, several observations reported by the authors suggest that referential processing in 3- and 5-year-old children is less geared toward the speaker's perspective than that in adults. For instance, Metzing and Brennan (2003) reported for adults a delay in both visually locating and touching an item, which was caused by the violation of a referential precedent by the partner who established it; they also reported a total absence of a delay in visually locating a referent associated with precedent violation by a new partner. In the Matthews et al. (2010) study, by contrast, children were significantly slower in their response to new labels even while interacting with a new partner. Moreover, the amount of children's verbal protests against the use of new labels was similar with the same or a new partner. Even when a new label was introduced by a new experimenter, some children corrected her in favor of the previously entrained term, which they claimed to better conform to the referent. (This behavior was clearly due to the experimental manipulation of labels, as both labels were welcomed by children during the training phase.)

Furthermore, delays associated with broken precedents were more important in the same partner condition than in the new partner condition only when the precedent was broken for the first time, with nonsignificant differences between conditions for the second trial. Matthews et al. (2010) suggested that the spillover effect of the first broken precedent in a first trial could have delayed children's processing of the next label item. The rationale is somehow complex here. For each participant, Matthews et al. (2010), submitted to the analysis a difference score, calculated as the subtraction of the time needed to find an item referred to with a maintained label from the time needed to locate another item referred to with a new label. The maintained label of a second trial directly followed the new label of the first trial. The authors suggest that the delay in processing the first new label could have affected the directly following maintained label of a second trial, thus annihilating potential differences with the use of a new label for the second time.

The invoked spillover effect put aside, the choice of analyzing a difference score for different items, instead of comparing entrained and broken labels for the same items, does not allow to distinguish two different partner-specific effects: the same speaker advantage in identifying referents for maintained descriptions and a new speaker advantage for broken precedents. As we mentioned previously, same speaker advantage in identifying maintained precedents may by entirely explained by low-level memory mechanism. For instance, it is possible that children better recognize a word repeated by the same speaker because a previously created association between this speaker's voice features and the word form has been stored in episodic memory. If so, facilitation in recognizing a word repeated by the same speaker would not require making inferences about the speaker's perspective. Memory traces, combining lexemes and phonetic features idiosyncratic to a speaker's manner of speaking may be activated independently from any assessments of common ground. By contrast, a potential advantage in identifying the referent for a broken precedent with a new speaker would likely result from consideration about a partner's perspective.

In the difference score used by Matthews et al. (2010), reaction times (RTs) for different items were collapsed into a single difference between one item with broken referential precedent and another with a maintained referential precedent. Matthews and colleagues thus subtracted absolute RT for an item designated by an entrained label item from the RT for another item designated by a new label. For example, the time needed to find a car in the test phase (designated as *car* in both parts of the experiment) was subtracted from the time to find a book in the test phase (designated as *story* in the entrainment phase and as *book* in the test phase). The fact that this difference score was found to be lower in the new partner condition can also be attributed to the priming of same label by same partner or to the expectation that a new partner is more likely to introduce new labels. In other words, this result is equally compatible with two different hypotheses, which are as follows:

1. Children were more delayed in reacting to broken precedent with the same versus a new speaker.

2. Children were less delayed in reacting to a maintained precedent with the same versus a new speaker.

However, it is crucial to disentangle these two interpretations, as only the first one would allow to reliably conclude that young children's reference resolution is guided by the assumptions about their partner's perspective.

Given the theoretical importance of children's spontaneous processing of conversational partner's perspective, we believe that it is important to carefully reexamine whether speaker's perspective is incorporated to the interpretation of referential expression in 3- and 5-year-old children. To this end, we conducted a conceptual replication of the study by Matthews et al. (2010). We implemented several methodological and analytical changes, detailed in the following text, in the spirit of refining this paradigm and gathering firmer evidence on referential processing in young children.

Method

Participants

To determine a sample size, we performed a simulation based on the estimates from the distribution (Gelman & Hill, 2007) reported in Matthews et al. (2010), fitted using lmer data function from the lme4 software package (Bates, Maechler, & Dai, 2008) for the difference score in the first trial. Power was estimated as the percentage of simulations that provide a *p* value smaller than 0.05. On the basis of this simulation, it should take 120 participants in both groups to reach 92% power for the condition effect (same vs. new partner); 80.3% power was estimated to be achieved with a sample of 60 children (30 in each group).

Our final sample included 3-year-old children ($n = 30$; $M = 3.6$ years; range = 3.2–4.1 years; 14 boys, 16 girls) and 5-year-old children ($n = 35$; $M = 5.6$ years; range = 4.8–6.5 years; 18 boys, 17 girls). The study protocol was approved by the Ethics Committee of the Queen Fabiola Children's University Hospital. Parents signed informed consent for their children prior to participation. Children in both groups were native French speakers, exposed to other languages less than 30% of the time. Four children in the 3-year-old group and one in the 5-year-old group were excluded because of technical problems with recordings, bilingualism issues, or experimenter's error.

Table 1 provides the descriptive statistics for both age groups. There was no difference between groups in socioeconomic background, as assessed by the Family Affluence Scale (Currie, Elton, Todd, & Platt, 1997), most families being middle class, or in nonverbal IQ, as assessed by the Leiter International Performance Scale–3 (Roid & Miller, 2013) with four intelligence subtests (Sequential Order, Form Completion, Classification and Analogies, and Figure Ground). In both groups, children's language skills were at least at the 3-year-old level, as assessed by the French version of the Peabody Picture Vocabulary Test–Revised (PPVT-R; Dunn, Thériault-Whalen, & Dunn, 1993). Unexpectedly, using age-normalized scores, children in the 3-year-old group scored significantly lower than those in the 5-year-old group. The French version of the PPVT-R was developed in Canada, and cultural differences between Belgian and Canadian French may explain this difference. Receptive vocabulary assessment can be influenced by cultural and linguistic dissimilarities in different geographical regions, and transferring a test on language abilities developed from one region to another is known to entail accuracy loss (Rondal, 1997; Thordardottir, 2014).

Table 1
Descriptive Statistics for the Two Age Groups

Characteristic	3 year olds ($n = 30$)	5 year olds ($n = 35$)	*t*	*p*
Gender (% female)	53.3	48.6		
Age *M* (*SD*)	3.62 (.23)	5.65 (.39)	–26.34	<.001
Socioeconomic background *M* (*SD*)	6.64 (1.39)	6.34 (1.52)	.064	.948
Nonverbal IQ *M* (*SD*)	99.45 (6.75)	99.03 (8.64)	.21	.832
Vocabulary *M* (*SD*)	101.71 (14.66)	112.83 (16.99)	–2.85	.006

Note. Children in the 3-year-old group ranged in age from 3.2 to 4.1 years; children in the 5-year-old group ranged in age from 4.8 to 6.4 years.

Stimuli

We created a collaborative task modifying the paradigm used in Matthews et al. (2010) in a way to simplify data collection and diminish potential coding biases. In the Matthews et al. study, participants rearranged toys in a block of pigeonholes. Video recording of the task were subsequently coded to extract the time intervals between experimenter's uttering the label and the child touching the corresponding object. Instead, we developed a computer task in which items could be moved around on a touchscreen display, which allows automatic recording of the precise time point at which the participant touches the target.

We used the Manulex (Lété, Sprenger-Charolles, & Colé, 2004) French database to select nouns that are frequently found in speech directed at 3-year-old children and are also actively used by first graders. Next, we narrowed down our choice to eight pairs of labels, such that each pair could constitute an alternative label for the same entity (see Table 2). Those entities could be referred to equally felicitously with two labels and thus corresponded to our four target items. A professional artist created two sets of 12 pictures, each consisting of the following:

- four pictures for target items, referred to with two alternative labels during the task;
- four pictures for nontarget items, always referred to with the same label; and
- four filler items, always referred to with the same label, but for which no measures were collected.

To ensure the same level of visual salience among items, drawings were created such that no picture stood apart in the amount of detail, hue, intensity, or color. Two sets of objects were created for a within-subject design, so that each child played with one set of objects during the first session and the other set on the second session (see the following text for details).

The task was created using Adobe Flash with ActionScript 2.0. The test screen consisted of four groups of three drawings

Table 2
French Labels Used in Set 1 and Set 2 (English Translations in Parentheses)

Target items	
docteur–médecin (doctor–physician)	enfant–bébé (child–baby)
peluche–nounours (cuddly toy–teddy bear)	fenêtre–vitre (window–pane)
feuille–papier (sheet–paper)	dame–femme (lady–woman)
chaussures–bottes (shoes–boots)	sac–cartable (bag–satchel)
Nontarget items	
avion (plane)	train (train)
souris (mouse)	lapin (rabbit)
cheval (horse)	téléphone (phone)
pomme (apple)	livre (book)
Filler	
poisson (fish), étoile (star), verre (glass), maison (house)	vache (cow), voiture (car), chien (dog), chat (cat)

in each corner of the screen and the drawing of a photo frame placed beneath a drawing of a photo camera in the center of the screen with the camera button above (see Figure 1). Exactly two items could be placed within the photo frame. To move an item to the photo frame, the child had to touch it first, then the item moved to the photo frame. The child's task was to make a set of photos following instructions by an experimenter sitting behind him or her. Once two items were placed in the picture frame, a photo could be made by touching the camera button. If the choice of two objects in the photo frame was correct, namely if it matched the experimenter's instructions, the sound of a photoflash was produced; if the choice was incorrect, an error sound was produced, and no photo could be taken. The script was run in Tobii studio software (Version 3.5.6; Tobii Technology, Inc., Stockholm, Sweden) as screen recording media element; all the clicks were recorded.

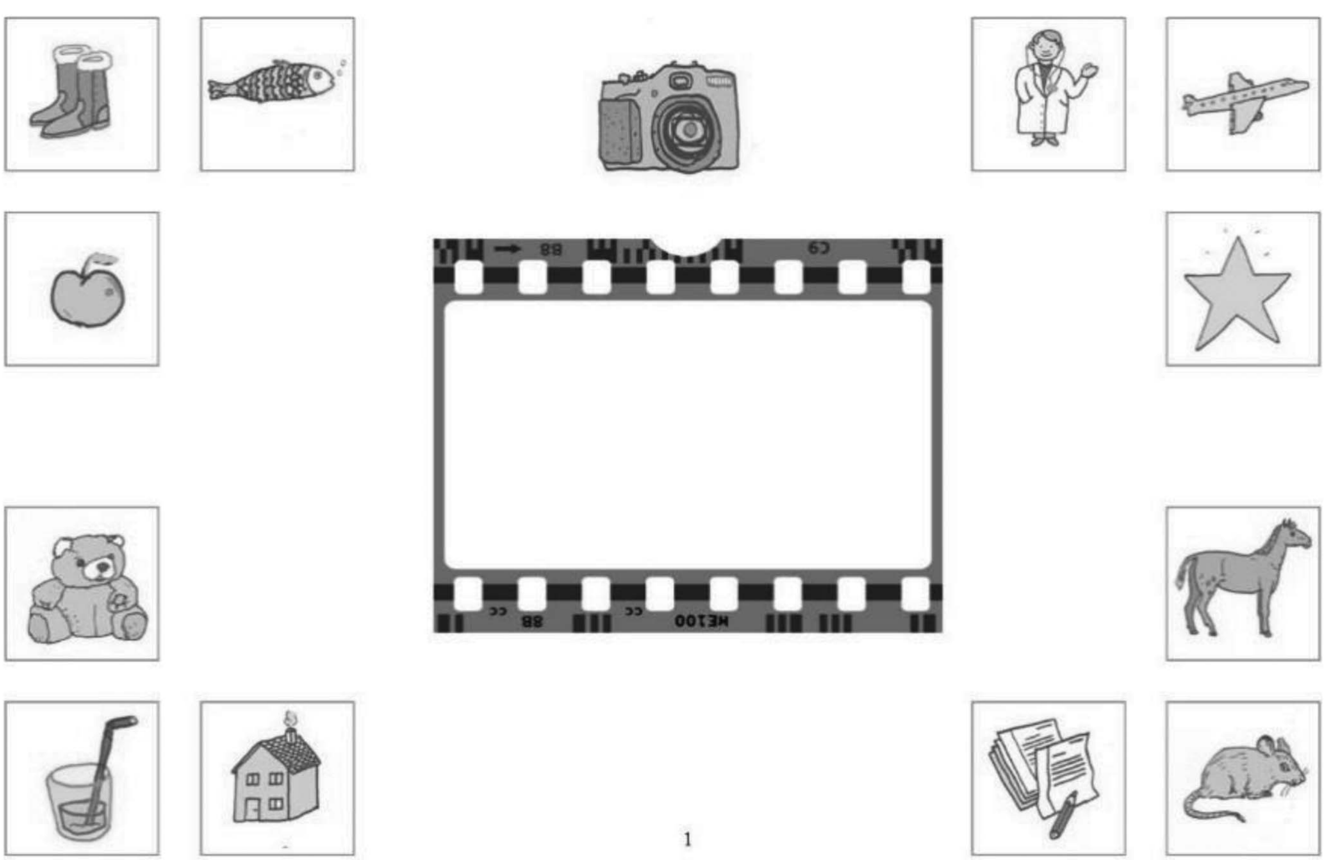

Figure 1. Screenshot of the starting screen for the Set 1. This figure is used with permission by Ostashchenko et. al.

Procedure

The experiment consisted of three sessions, each lasting around 30 min. During the first session, two experimenters were introduced to each child. One experimenter administered the PPVT-R and two of the four subtests of the Leiter-R, so that both experimenters were familiar to the child at the end of the first session.

In the second session, the child saw again the same experimenters. The session began by the administration of the remaining two subsets of the Leiter-R (by the same experimenter as in the first session). Next, after some time of free play with both experimenters, one experimenter left the room with the phone of the other experimenter. The experimenter who stayed in the room showed the child the first pictures of a set of photos that she had on an electronic device and explained that the same photoset could be reproduced on the touchscreen display. After showing how to make photos on the screen, the experimenter asked if the child would like to play. When the child agreed, the first part of the experiment started. The experimenter began by telling the child the following: "Now, we need to take a photo of the [first item], can you touch it?" Once the child touched the first item, the experimenter went on to say the following: "And now we need [second item], can you touch it?" Finally, when both items were located within the photo frame, the experimenter said, "And now we can take a photo, touch the camera!"

A substantial advantage of our instructions is that they do not contain any spatial prepositions (e.g., *next to*, *under*, or *above*). By contrast, the nature of the Matthews et al. (2010) design forced them to spatially define the place to which children had to move the items. In this way, we avoid potential errors and delays linked to cognitive processing of spatial relations.

During the first entrainment part of the task session, the child made 16 photos reflecting different combinations of pairs of items. Those combinations were created so that each target and nontarget item was referred to with the same expression three times; the remaining eight items were randomly selected from the list of fillers. Experimenters were instructed to pronounce the second label word at the exact moment when the previous item has been placed in the photo frame for the accuracy of collected RTs.

Once 16 photos were made, the first part of the session was over and the second experimenter came back. She explained that the first experimenter received several phone calls and probably should answer her phone. What happened next depended on the condition—new partner or same partner—that was assigned to the set. In the new partner condition, the first experimenter replied that she was waiting for important phone calls and should answer her phone. She asked whether the child would like to continue making photos with the second experimenter and left the room after the child's approval. In the same partner condition, the first experimenter answered that she would like to call back later and continued to play with the child.

The second part of the session began next, with the same or the new experimenter—depending on the condition—playing the same photo game. Eight photos in total were made in the second session. For each photo, the first item referred to was a filler, and the second alternated between target and nontarget items. All target items were designated by the label different from the one used in the first training session (i.e., new label). In this way, each target and nontarget object was always followed by a filler. This allowed us to eliminate any potential spillover effect between target and nontarget items. The location of the items on the screen in each set was kept constant throughout all trials to minimize the potential influence of individual variation in executive processing and visual and/or spatial perception on RTs.

In this within-subject design, each child took part in two experimental sessions, one per condition (same vs. different partner), each session consisting of two parts (entrainment and test phase). Each experimental session was associated with one of the two sets of pictures, and with one of two conditions. The association between sets of pictures and conditions was evenly counterbalanced across participants in both groups (3- and 5-year-old children), so that for half of the children the first session consisted of the new experimenter condition, whereas the other half started with the same experimenter condition. Experimental sessions corresponding to two conditions were separated by at least 2 days of interval; in 90% of cases, an interval between session exceeded 1 week (range = 2 to 28 days; $M = 8.7$, $SD = 6$). All experimenters had native or native-like fluency in French. Experimenters were counterbalanced across participants so that each member of experimenter pair was engaged in the same partner condition an equal number of times.

It could be that some children are "hyper-conventional" in the sense that they would perceive one of the two labels associated with a target as an intrinsically inaccurate description, independently of the rupture of lexical entrainment. To reduce this potential bias, labels for target items were also counterbalanced between the first and second parts of each set. For example, for half of the children assigned to Set 1, the picture of the physician in a white coat was referred to by *médecin* in the first part of the session and by *docteur* in the second part; this order was reversed for the second half of children assigned to Set 1.

Recall also that in Matthews et al. (2010), the effect of the referential pact was only significant for the first trial. To ensure that any difference between trials was not due to the peculiarities of some items, we also counterbalanced the order of the four target items in the second part of session. Using a Latin-square design, we created four scripts for each set, corresponding to different orders of appearance of target items. This modification also allowed to reduce potential item effects that could interfere with the investigated referential processing effects.

Our experimental design is thus identical in crucial aspects to the paradigms used by Matthews et al. (2010) or, for that matter, by Metzing and Brennan (2003). However, we modified their task in the three following aspects:

- touchscreen display to facilitate and consolidate the procedure of RT collection;
- four, instead of two, critical trials to further examine the trial order effect reported in Matthews et al. (2010); and
- unambiguous instructions with no spatial prepositions to reduce errors independent of referential processing.

Results

Data Preparation

Our focus is on RTs, understood as the time needed to locate an item referred to by the experimenter. RTs were obtained by manually segmenting all screen recordings obtained in Tobii studio software. To generate these segments, we identified the onset of each critical word through visual and audio inspection of the corresponding sound waves in PRAAT (Boersma, 2001). Next, we tagged the time points corresponding to this onset on screen recordings for all target and nontarget items. Finally, we isolated all the segments between those tags and the corresponding clicks on the screen to analyze their length.

Importantly, in addition to the absolute RT values (in ms) for the second part of each session, we also computed relative RTs for each item, by subtracting the length of the corresponding segment in the first part of the session from the length of the corresponding segment in the second part. As an illustration, take again the target item representing a physician in white coat, which in French can be referred to as *docteur* or *médecin*. If *docteur* was used during entrainment phase and *médecin* after the interruption, we subtracted the absolute RT for the last use of *docteur* during entrainment phase from the absolute RT interval for *médecin*. We checked the distribution of both relative and absolute RTs. Unsurprisingly, absolute RTs were positively skewed; these values were consequently log-transformed using the Box–Cox procedure (Box & Cox, 1964).

Means of absolute RTs are reported in the Table 3 as a function of condition (same vs. new partner) and referential term (maintained vs. broken precedent) for both ages in four trials.

Analyses

Accuracy in finding the item referred to by the experimenter was high in both age groups (95% for the 3-year-old group and 99% for the 5-year-old group). We excluded from the analysis trials in which the experimenter used the wrong label (0.7%), the child touched the wrong item (3%), the child asked for help and exceeded the 10-s interval of self-reliant searching (1.5%), or the child asked questions not related to the task (0.3%). The remaining 924 trials for relative RTs and 969 trials for absolute RTs in the test phase were submitted to analysis.

Relative RTs and log-transformed absolute RTs were analyzed separately using linear mixed-effect models with by-item and by-participant intercepts in the random structure. The significance of a factor was analyzed by building it as fixed factor within a model and conducting log-likelihood comparisons of fit with otherwise identical models which excluded this factor (Baayen, Davidson, & Bates, 2008). The following four factors were analyzed: referential precedent (maintained vs. broken), condition (new vs. same partner), trial (1 to 4), and group (3- vs. 5-year-old children). All analyses were implemented with the lme4 software package (Bates et al., 2008) in R (R Core Team, 2017). Reported effect size for each factor corresponds to a partial eta squared for this factor that was calculated by working backward from the *F* values output by SAS PROC MIXED (Selya, Rose, Dierker, Hedeker, & Mermelstein, 2012).

Starting with relative RT values, there was a significant effect of referential precedent, $\chi^2(1) = 24.194$, $p < .001$, $\eta_p^2 = 0.04$. However, there were no effects of condition, $\chi^2(1) = 1.127$, $p = .288$, $\eta_p^2 < 0.001$, trial, $\chi^2(3) = 2.869$, $p = .09$, $\eta_p^2 = 0.003$, and group, $\chi^2(1) = 0.64$, $p = .423$, $\eta_p^2 < 0.00$). In sum, relative RTs were significantly higher for the entrained referential precedent was broken versus maintained (see Table 4 for a summary of the best fitting model). Figure 2 contains boxplots depicting distribution of relative RT values for maintained and broken precedents in the two groups. Summary of the statistical output model specifying random and fixed effects for relative RTs are provided in Table 4.

Turning to absolute RTs, there was also a significant effect of referential precedent, $\chi^2(1) = 6.544$, $p = .01$, $\eta_p^2 = 0.011$. Here, however, we also found significant effects of group, $\chi^2(1) = 31.146$, $p < .001$, $\eta_p^2 = 0.036$, but no interaction between referential precedent and group, $\chi^2(1) = 0.492$, $p = .482$, $\eta_p^2 < 0.001$. There was also a significant effect of condition, $\chi^2(1) = 4.668$, $p = .03$, $\eta_p^2 = 0.004$, but no interaction between condition and referential precedent, $\chi^2(1) = 1.375$, $p = .242$, $\eta_p^2 < 0.001$, and between condition and group, $\chi^2(1) = 0.01$, $p = .927$, $\eta_p^2 < 0.001$. Finally, there was no effect of trial, $\chi^2(1) = 2.08$, $p = .149$, $\eta_p^2 < 0.001$. As is shown in Table 5, which summarizes the final model, absolute RTs of 3-year-old children were longer across all items and trials, in both the first and second part of experimental sessions. Although 3-year-old children were overall slower to respond, both groups were delayed by the rupture of referential pacts (independently of the condition) and by a new partner (independently of whether the referential pact was maintained or broken).

Discussion

Both relative and absolute RTs show that preschoolers are significantly delayed when a previously entrained label is replaced

Table 3
Mean Reaction Times (in s; Standard Deviations in Parentheses) to Touch Nontarget (Maintained Precedent) and Target (Broken Precedent) Items

	Trial 1		Trial 2		Trial 3		Trial 4	
Age group/condition	Maintained	Broken	Maintained	Broken	Maintained	Broken	Maintained	Broken
3 year olds								
Same partner	2.54 (.82)	3.04 (1.80)	2.96 (1.58)	2.74 (1.70)	2.60 (.80)	2.98 (1.80)	2.30 (.80)	2.46 (1.02)
New partner	2.76 (1.40)	3.50 (2.28)	2.82 (.98)	3.76 (1.70)	2.40 (.64)	3.80 (2.16)	2.48 (.92)	2.36 (1.02)
5 year olds								
Same partner	1.56 (.50)	2.12 (1.04)	1.86 (.72)	2.40 (1.98)	2.08 (.90)	2.22 (.84)	2.02 (1.50)	2.14 (1.06)
New partner	1.82 (.72)	2.28 (1.22)	2.20 (1.34)	2.64 (1.80)	2.04 (1.10)	2.00 (.76)	2.08 (.76)	2.38 (1.52)

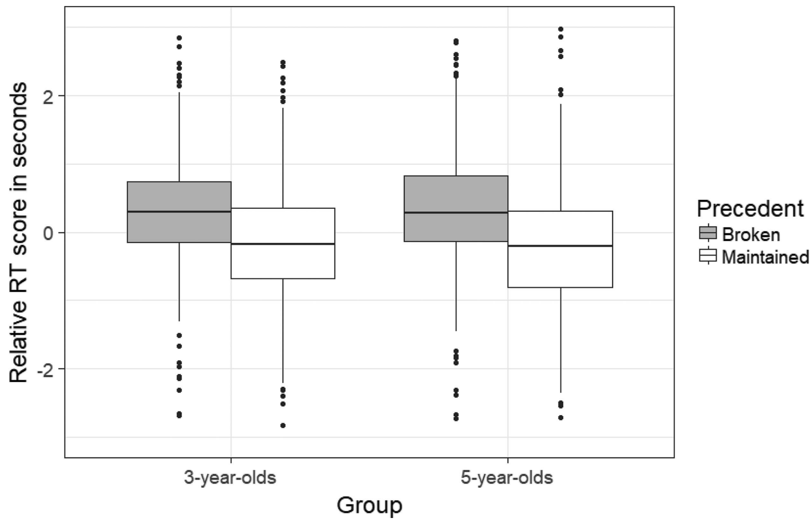

Figure 2. Boxplots of relative reaction time (RT) scores for broken and maintained precedents.

by another one. The comparison of these two different RT measures shows that this effect is age-independent. Although absolute RTs were higher for 3-year-old children on any kind of item, there was no interaction between group and referential precedents on relative RTs. In other words, although 3-year-old children were generally slower in their processing of verbal instructions or in their sensory-motor responses, there was no difference between ages in the delay due to violation of lexical entrainment. Another crucial result is that— contrary to the predictions of the partner specific view—this delay does not decrease when the new label is used by a new speaker who was absent during the interaction that resulted in the entrainment of the first label.

This is not to say that children were insensitive to the identity of their interactional partner. Recall that there was a significant condition (new vs. same speaker) effect on absolute RTs, suggesting that switching to a new interactional partner entailed a delay for previously entrained and new labels alike. In addition to lexical entrainment, our results evidence, so to speak, entrainment for the conversational partner. The advantage in matching the description with its referent emerging in the same partner condition is consistent with Graham et al. (2014), albeit in their study the effect was observed only for the previously entrained description. To repeat, however, the processing advantage linked to a familiar partner does not demonstrate integration of common ground information to referential processing.

Our results also clearly show that partner's perspective is not yet fully integrated within referential processing in young children. This conclusion partly contradicts that by Matthews et al. (2010), who reported limited evidence for a partner-specific referential processing in 3-year-old children. Our procedure was not completely identical to the one used by these authors. However, the three major modifications we made (i.e., use of touchscreen instead of real shelf, four instead of two critical trials, and exclusion of spatial preposition from instructions) should not have affected children's sensitivity to their partner's perspective, were it inherent in the precedents processing. Even though in the present study the task was displayed on a touchscreen, all of the instructions were

Table 4
Summary of the Output Model Specifying Random and Fixed Effects of a Mixed Linear Model Regressing Relative Reaction Times on Broken Precedent

Effects	Estimate	*SE*	*t*	Variance	*SD*
Fixed					
Intercept	202.77	78.55	–2.581		
Broken precedent	762.85	106.99	7.130		
Random					
Participant					
Intercept				6,284	79.27
Item					
Intercept				32,079	179.11

Note. $N = 65$; Number of observations = 924; Number of items = 24.

Table 5
Summary of the Output Model Specifying Random and Fixed Effects of a Mixed Linear Model Regressing Log-Transformed Absolute Reaction Times on Group, Referential Precedent, and Condition

Effects	Estimate	*SE*	*t*	Variance	*SD*
Fixed					
Intercept	7.813	.051	153.71		
5 year olds	–.295	.046	–6.32		
Broken precedent	.135	.047	2.85		
New partner	.055	.024	–2.22		
Random					
Participant					
Intercept				.026	.163
Item					
Intercept				.010	.101

Note. $N = 65$. Number of observations = 969. Number of items = 24.

produced during a real-time interaction so as to take into account the higher likelihood to observe a partner-specific effect in interactive settings rather than with prerecorded instructions (Brown-Schmidt, 2009a). Furthermore, partner change was always made salient to the child in the same way as in the Matthews et al. study.

One could wonder whether visual representations of the items on a screen instead of real objects could have influenced children's expectations about a new partner's choice of referential expressions. Although we cannot fully rule out this possibility, we could not find any theoretical motivation for it. By contrast, our computerized method carries several methodological benefits. To begin with, in simple word-recognition tasks used with 1- to 4-year-old children, the touchscreen method has been found to yield reliable RT and accuracy data compared with eye-tracking and in-person storybook paradigms (Frank, Sugarman, Horowitz, Lewis, & Yurovsky, 2016; Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015). The use of a touchscreen in the present study also allowed to accurately record children's haptic responses. In the Matthews et al. (2010) design, children were trained to place their hands on stickers attached to the surface of the table to ensure subsequent RT coding; this part of the instructions could have put additional cognitive demand on the children who have not yet acquired reliable motor control over their hand movements at the age of 3. In addition to avoiding spatial prepositions (see the preceding text), the use of the touchscreen also allowed us to implement a feedback that could positively reinforce children's motivation to react in an accurate and rapid way to the experimenter's instructions.

The discrepancy between our results and those of Matthews et al. (2010) results may boil down to a difference in analytical strategy. Our goal was to inspect two different mechanisms—namely priming of the maintained precedent by the original speaker and integration of a new partner's perspective—that potentially explain partner effects in lexical entrainment tasks. Consequently, we chose not to concatenate RTs for different items. Matthews et al. (2010), on the contrary, transformed their absolute RTs into difference reaction scores for items of different type. As explained in the introductory paragraphs of this article, this manipulation could have obfuscated the potential contrast between two different mechanisms, potentially involved in referential processing. By assessing both relative RTs—constructed as the difference between response to a given item in the first and the second parts of the experimental session—and (log-transformed) absolute RTs for all items in the second part of a session, we made sure to keep correct by-item variance. In this way, we were also able to analyze label effect as a fixed factor and to distinguish the effect related to the use of a low-level priming mechanism from potential effects of expectations about partner's perspective. The results obtained for the two transformed data sets were consistent, with the second method being more sensitive to the small fixed factors effects.

As expected from the literature, though the effect of broken precedents clearly emerged from our results, we could not detect the effect of partner going in the direction predicted by Matthews et al. (2010; or at least in accordance with one of the possible explanations put forth by those authors). If anything, the modest partner effect we observed with our sample of 65 children went in the direction opposite to that previously described in adults. Contrary to the predictions of the partner-specific view, we found that the introduction of a new partner was associated with longer RTs, both with entrained and violated referential choices.

Conclusion

One way to determine whether children spontaneously and by-default adopt their conversational partner's perspective during lexical processing is to investigate whether their expectation about referential descriptions are consistent or not with their partner's perspective. The use of a new description by the partner with whom one has previously entrained a different description of the same object is conversationally inconsistent; by contrast, a new partner cannot be expected to abide by a referential precedent of which she has no knowledge. According to the results published by Matthews et al. (2010), children were particularly delayed reacting to a broken precedent compared with a maintained one, this difference being dependent on the partner condition in which it occurred. We performed a conceptual replication of this study, aiming at disentangling two possible explanations of this effect. The advantage in locating an item referred to with a maintained label when interacting with the same partner would indicate children's reliance on automatic priming during referential processing. Shorter delays in finding an item referred to with a new label in a new partner condition would provide evidence that young children have already integrated their partner's perspective within their processing of referential expressions. Using a robust by-item and by-participant analysis, we found that reactions to the rupture of referential precedents do not depend on the partner's identity. Our results also indicate that lexical processing in young children is facilitated by previous interaction with their conversational partner. Both familiar and novel labels were processed a little faster by the children in both age groups when the conversational partner did not change (the approximate gain was about 5%). This advantage likely reflects children's adapting to a partner through the task, resulting in more efficient language processing. With a new partner, the adaptation process was disrupted, taxing referential processing.

Several studies suggest that processing of relevant lexical information might be facilitated by partner-specific memory association. One's conversational partner identity is a strong contextual cue, which may influence memory retrieval (Barr, Jackson, & Phillips, 2014; Horton & Gerrig, 2005). Our finding that children adapt to their partner—not only to find referents for previously used descriptions, but also to match new labels with available referents—raises interesting questions about the level of information cued by the partner. According to partner-specific accounts, people are cued as unique perspectives they offer (Brown-Schmidt, Yoon, & Ryskin, 2015). However, our results suggest that, at least in young children, interactional partners are cued via low-level priming mechanisms as the sources of phonological information, inducing more efficient speech processing (Shintel & Keysar, 2009).

Our results show early egocentric processing in comprehension of referential descriptions. Even though young children cannot yet conceive of their partners in terms of perspectives that they offer, children do benefit from low-level adapting mechanisms during conversations. Consistent with previous studies (Doherty & Perner, 1998; Graham et al., 2014; Matthews et al., 2010; Perner, Stummer, Sprung, & Doherty, 2002), we found that violations of lexical entrainment significantly delay children's processing. Unlike adults, who can easily follow modifications in lexical choices of their conversation partners, children display difficulties in disengaging from the entrained and therefore more salient conceptual perspective. Maturation of executive function skills may provide children with more flexible use of conceptual perspective that, in turn, would allow further integration of common ground strategy within referential processing.

It remains to be seen whether potential sensitivity to a partner's perspective could be elicited in a task where descriptions are more closely associated with a particular partner. Like Matthews et al. (2010), we used conventional noun descriptions for objects. However, the major part of the literature on referential precedents in adults uses nonce denominations of abstract figures; such descriptions (e.g., *shiny cylinder* and *silver tube*) are arguably more dependent on the partner who introduces them. In adults, effects of referential precedents are less important for conventional nouns (Barr & Keysar, 2002). One could still speculate, however, that in children the use of alternative conventional nouns does introduce a greater shift in perspective because they display stronger preference for a conventional referring expression once it has been evoked. It could then be possible that this conflict obfuscates any potential partner-specific effects in referential processing. If so, other nonconventional referential expressions could increase the likelihood of detecting encoding a partner's perspective in children. That said, in the study by Graham et al. (2014), where an adjectival modifier was used to create alternative expressions, the only reported advantage was related to the maintained precedent with the same partner.

Ongoing debates on the cognitive mechanisms that underpin coordination of meanings between a listener and a speaker may benefit from investigating partner-related cues during referential processing in children. The literature on adults suggests the existence of a robust expectation that speakers be consistent in their referential choices. Evidence of such partner-sensitive expectations early in childhood would have constituted a strong argument in favor of a view that common ground is inherent in utterance interpretation. However, our results show that access to meaning in 3- and 5-year-old children is not modulated by expectations about a partner's perspective in an adult-like way.

References

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. http://dx.doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Jackson, L., & Phillips, I. (2014). Using a voice to put a name to a face: The psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General, 143,* 404–413. http://dx.doi.org/10.1037/a0031813

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language, 46,* 391–418. http://dx.doi.org/10.1006/jmla.2001.2815

Bates, D., Maechler, M., & Dai, B. (2008). *The lme4 package*.

Boersma, P. (2001). PRAAT: A system for doing phonetics by computer. *Glot International, 5,* 341–345. http://dx.doi.org/10.1017/CBO9781107415324.004

Box, G., & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26,* 211–252.

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

Brown-Schmidt, S. (2009a). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language, 61,* 171–190. http://dx.doi.org/10.1016/j.jml.2009.04.003

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

Brown-Schmidt, S., Yoon, S. O., & Ryskin, R. A. (2015). People as contexts in conversation. *Psychology of Learning and Motivation, 62,* 59–99. http://dx.doi.org/10.1016/bs.plm.2014.09.003

Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 591–610. http://dx.doi.org/10.1037/xlm0000345

Currie, C. E., Elton, R. A., Todd, J., & Platt, S. (1997). Indicators of socioeconomic status for adolescents: The WHO health behaviour in school-aged children survey. *Health Education Research, 12,* 385–397. http://dx.doi.org/10.1093/her/12.3.385

Deliens, G., Antoniou, K., Clin, E., & Kissine, M. (2017). Perspective-taking and frugal strategies: Evidence from sarcasm detection. *Journal of Pragmatics, 119,* 33–45. http://dx.doi.org/10.1016/j.pragma.2017.08.002

Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development, 13,* 279–305. http://dx.doi.org/10.1016/S0885-2014(98)90012-0

Dunn, L., Thériault-Whalen, C., & Dunn, L. (1993). *Echelle de vocabulaire en images Peabody: Adaptation Française du Peabody Picture Vocabulary Test–Revised*. Toronto, Ontario, Canada: PSYCAN.

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

Frank, M. C., Sugarman, E., Horowitz, A. C., Lewis, M. L., & Yurovsky, D. (2016). Using tablets to collect data from young children. *Journal of Cognition and Development, 17,* 1–17. http://dx.doi.org/10.1080/15248372.2015.1061528

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

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27,* 181–218. http://dx.doi.org/10.1016/0010-0277(87)90018-7

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, UK: Cambridge University Press.

Graham, S. A., Sedivy, J., & Khu, M. (2014). That's not what you said earlier: Preschoolers expect partners to be referentially consistent. *Journal of Child Language, 41,* 34–50. http://dx.doi.org/10.1017/S0305000912000530

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

Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition, 149,* 104–120. http://dx.doi.org/10.1016/j.cognition.2015.12.008

Hendrickson, K., Mitsven, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2015). Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental Science, 18,* 723–735. http://dx.doi.org/10.1111/desc.12250

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

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11,* 32–38. http://dx.doi.org/10.1111/1467-9280.00211

Kronmüller, E., & Barr, D. J. (2006). Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language, 56,* 436–455. http://dx.doi.org/10.1016/j.jml.2006.05.002

Kronmüller, E., & Barr, D. J. (2015). Referential precedents in spoken language comprehension: A review and meta-analysis. *Journal of Memory and Language, 83,* 1–19. http://dx.doi.org/10.1016/j.jml.2015.03.008

Kronmüller, E., Noveck, I., Rivera, N., Jaume-Guazzini, F., & Barr, D. (2017). The positive side of a negative reference: The delay between linguistic processing and common ground. *Royal Society Open Science, 4,* 160827. http://dx.doi.org/10.1098/rsos.160827

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers, 36,* 156–166. http://dx.doi.org/10.3758/BF03195560

Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing how rabbits quack and competitors act: Limits on preschoolers' efficient ability to track perspective. *Child Development, 85,* 1519–1534. http://dx.doi.org/10.1111/cdev.12224

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

Matthews, D., Lieven, E., & Tomasello, M. (2010). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology, 46,* 749–760. http://dx.doi.org/10.1037/a0019657

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language, 49,* 201–213. http://dx.doi.org/10.1016/S0749-596X(03)00028-7

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308,* 255–258. http://dx.doi.org/10.1126/science.1107621

Pechmann, T., & Deutsch, W. (1982). The development of verbal and nonverbal devices for reference. *Journal of Experimental Child Psychology, 34,* 330–341. http://dx.doi.org/10.1016/0022-0965(82)90050-9

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development, 17*(3–4), 1451–1472. http://dx.doi.org/10.1016/S0885-2014(02)00127-2

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

R Core Team. (2017). *R: A language and environment for statistical computing*. Retrieved from https://www.r-project.org/

Roid, G. H., & Miller, L. J. (2013). *Leiter International Performance Scale* (3rd ed.). Wood Dale, IL: Stoelting.

Rondal, J.-A. (1997). *L'evaluation du langage*. Brussels, Belgium: P. Mardaga.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1255–1266. http://dx.doi.org/10.1037/a0018729

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General, 141,* 433–438. http://dx.doi.org/10.1037/a0025458

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology, 3,* 111.

Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. *Topics in Cognitive Science, 1,* 260–273. http://dx.doi.org/10.1111/j.1756-8765.2009.01018.x

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18,* 587–592. http://dx.doi.org/10.1111/j.1467-9280.2007.01944.x

Thordardottir, E. (2014). The typical development of simultaneous bilinguals: Vocabulary, morphosyntax and language processing in two age groups of Montreal preschoolers. In T. Grüter & J. Paradis, *Input and experience in bilingual development* (pp. 141–160). Amsterdam, the Netherlands. John Benjamins Publishing Company. http://dx.doi.org/10.1075/tilar.13.08tho

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Whitehurst, G. J., & Sonnenschein, S. (1978). The development of communication: Attribute variation leads to contrast failure. *Journal of Experimental Child Psychology, 25,* 490–504. http://dx.doi.org/10.1016/0022-0965(78)90071-1

Received January 29, 2018
Revision received June 27, 2018
Accepted June 29, 2018 ■