Accepted manuscript (post-peer review, pre-copyedit), *Perspectives on Psychological Science*.
Please do not cite this version.

A New Look at Young Children's Referential Informativeness

Jared Vasil[1]

[1] Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

Keywords:

children, common ground, active inference, reference production

Word counts:

Main Text: 14,665

Main Text + Endnote + Figure Captions: 15,483

Accepted manuscript (post-peer review, pre-copyedit), *Perspectives on Psychological Science*.
Please do not cite this version.

Abstract: This article reviews experimental evidence for the dependence of 2- to 5-year-olds' linguistic referential informativeness on cues to common ground (CG) and proposes a process model. Cues to CG provide evidence for CG, that is, for the shared knowledge, beliefs, and attitudes of interlocutors. The presence of cues to CG (e.g., unimpeded listener line of regard or prior mention) is shown to be associated with less informative reference (e.g., pronouns). In contrast, the absence of cues to CG (e.g., impeded listener line of regard or new mention) is shown to be associated with more informative reference (e.g., nouns). Interestingly, informativeness is sensitive to linguistic before nonlinguistic cues to CG (i.e., 2.0 vs. 2.5 years of age, respectively). Reference is cast as a process of active inference, a formulation of Bayesian belief-guided control in biological systems. Child speakers are hierarchical generative models that, characteristically, expect sensory evidence for the evolved, prior Bayesian belief that interlocutor mental states are aligned (i.e., that CG exists). Referential control emerges as an embodied tool to gather evidence for this prior belief. Bottom-up cues to CG elicited by action drive updates to beliefs about CG. In turn, beliefs about CG guide efficient referential control.

Introduction

Among primates, human communication is unique in that it is characteristically cooperative. Speakers rely on the knowledge, beliefs, and attitudes they share with listeners – that is, their *common ground* (CG) – to package their communicative intentions in ways that listeners can readily interpret (Clark, 1996; Tomasello, 2008). In doing this, speakers frequently produce linguistic *referring expressions* with the intention of inviting listeners to jointly attend to *intended referents* (i.e., the entity to which speakers wish to jointly attend; Tomasello, 1998). Speakers' choice of referring expression depends on the inferred CG status of intended referents. Speakers typically use less informative expressions, like pronouns (e.g., *it*), to refer to referents in CG. Conversely, speakers tend to use more informative expressions, like lexical noun phrases (NPs, e.g., *the cup*) when placing referents into CG. That is, the relationship between CG and reference is an inverse one: More CG is associated with less informative reference, while less CG is associated with more informative reference. The upshot is that using CG saves on communicative cost: Speakers use the cheapest referring expressions they can "get away with," given their inferences about CG (Ariel, 1988; Gundel et al., 1993; see Bohn & Köymen, 2018).

Being a ubiquitous and uniquely human trait, it is especially interesting to ask: How do we come to master referential communication? Historically, the approach taken by researchers who have asked this question has been to evaluate children's reference using adult communication as yardstick. This approach has resulted in what may be called the "deficits" view of early referential communication. A classic discussion of children's reference illustrates the deficits view (Glucksberg et al., 1975). Glucksberg et al. (1975) argued that – while "egocentrism and role-taking" (p. 324) are unhelpful concepts for analyzing early reference – children's "speaking… skills appear to improve with age, with nearly adult competence attained by early teens," (p. 336).

To support these claims, Glucksberg et al. (1975) reviewed literature on children's production in referential communication tasks. In these tasks, children single out, via reference, an intended referent from nonreferents – that, together, comprise a *referential array* – following manipulations to (cues to) the CG status of referents. Thus, in discussing the results of Glucksberg et al. (1966), the authors write that "dyads were incapable of adequate communicative performance" (p. 317) before stating that "young children communicate poorly" as a result of "deficiencies in the… skills involved [in referential communication]," (both p. 332). The deficits view thus advances the fundamentally negative point that young children are, essentially, bad at reference and that it is not until "their early teens [that children demonstrate] normal adult competence," (p. 319).

The deficits view is misguided and misses what is interesting about referential communication. Selection favors traits that enhance fitness within life history stages; natural selection is blind to teleological adult endpoints (Bjorklund & Ellis, 2014). Consequently, if the goal is to characterize the ontogeny of a uniquely human trait, like referential communication, then it is more theoretically enriching to know whether, how, and which experiences influence its expression in ontogeny than it is to know its age of mature expression in the phenotype (Tomasello, 2019).

Against the deficits view, we argue for a positive view on early reference, one which better addresses the question that began the previous paragraph. On this view, fore-fronted is the ability – possessed by young children and adults alike – to scale linguistic informativeness according to inferences about CG. In support of this argument, experimental research is reviewed that suggests that children use inferences about CG to appropriately scale their linguistic informativeness by 2;0 (years;months). By adopting the positive view on early reference, one must conclude that even young children have "adult competence" – a far cry from the suggestion of the deficits account.

This conclusion results from shifting the definition of adult competence from 'doing the same thing as adults' (deficits view) to 'doing the same *kind* of thing as adults' (positive view), that is, scaling linguistic informativeness inversely with CG (as adults do in referential communication tasks; e.g., Winters et al., 2018).

The positive view is motivated by a model of human communication as active inference (Vasil et al., 2020). This model portrays speakers as hierarchical models of the unobservable causes of sensation (i.e., of the minds of agents "like me"). Communicative action is cast as motivated, cost-sensitive control guided by the prior (Bayesian) belief that one's mental states are statistically similar to those of conspecifics. Referential production emerges in these models as a characteristic means to gather sensory evidence for this prior belief, on average across time. This means that children are hypothesized (i) to be characteristically motivated to communicate to align mental states with others (e.g., via reference); and (ii) to rely on *cues to common ground* to guide their expenditure of communicative effort (i.e., their informativeness). For example, if a single cup were to comprise a jointly attended referential array, then children might use this nonlinguistic cue to CG to infer that they may felicitously use *it* to refer to the cup, while no such inference would be licensed were there multiple cups (instead, *the red cup* would be needed; but see Clark et al., 1983). Similarly, if a red cup were discourse topical, then children could use this linguistic cue to CG to infer the felicity of endophoric coreference using *it*, while *the red cup* would be needed for contrastive reference. In short, Vasil et al. claim that – depending on how cues to CG influence their inferences about CG – children might feel sufficiently confident, or not, to refer using *it* over *the red cup*. Specifically, children are argued to perceive cues to CG as evidence for probabilistic hypotheses about CG (relatedly, Bohn & Frank, 2019) and to use these hypotheses to calibrate their informativeness following the inverse pattern, above.

There has thus far been no review of the evidence for a general principle for early referential production like that advanced here, namely, the inverse relationship between CG and referential informativeness (e.g., Ateş-Sen & Küntay, 2015; de Cat, 2015; Graf & Davies, 2014; Nilsen & Fecica, 2011). The utility of this principle is demonstrated in a review of a large body of literature. Two points about that literature. First, because we are concerned with whether and how informativeness tracks with cues to CG, this article reviews literature that operationalized informativeness in terms of the forms children used (e.g., pronouns, lexical NPs), not the "optimality" of their informativeness (e.g., Davies & Katsos, 2010). Second, understanding how cues to CG influence referential production amounts, in part, to understanding the causal influence of the sensory "microstructure" (Frank et al., 2013, p. 21) of referring situations on children's inferences about CG. To understand this microstructure, this review examines subtle methodological details to critically discuss and compare studies (Graf & Davies, 2014). The review is organized in halves. The second section reviews research on nonlinguistic cues to CG and the third section research on linguistic cues to CG. Figure 1 provides a roadmap for and summary of the review. A fourth section elaborates the model that motivates the positive view (Vasil et al., 2020)

Nonlinguistic Cues to Common Ground and Young Children's Referential Informativeness

This section discusses nonlinguistic cues to CG. We label conditions in the reviewed literature that included more cues to CG the "more CG" and those that included fewer cues to CG "less CG" conditions (with variants as needed). This section is split into three subsections. Each subsection discusses literature on one type of nonlinguistic cue to CG (listener line of regard, array

configuration, and nonlinguistic cues to cultural CG) and begins with an introductory paragraph followed by a general discussion, discussion of null results (where applicable), and a conclusion.

Listener Line of Regard

Beginning around 9 to 12 months of age, infants begin to jointly attend towards triadic entities with caregivers (Carpenter et al., 1998). For infants and young children, a key behavioral marker of joint attention is the "look back," a listener gaze alternation between child and triadic entity (Carpenter & Liebal, 2011). Observing look backs ensures infants that intended referents are the focus of joint attention and, thus, in CG with listeners (Liszkowski et al., 2004, 2007). Importantly, "genuine" look backs require that listeners have unimpeded lines of regard towards intended referents. For instance, the presence of barriers that block listener visual access to intended referents precludes the possibility of genuine look backs. In such cases, there cannot be (visual) joint attention to, and therefore no CG about, referents. For adult speakers, one way to nonetheless ensure referents are placed into CG is to produce more informative reference when listener visual access is impeded compared to unimpeded. When do children follow this pattern?

General discussion. In an early investigation, O'Neill (1996) examined whether 2-year-olds adjust their informativeness according to listener line of regard. In O'Neill (1996, Study 1), children (2;7) requested a toy from caregivers. Children watched as an experimenter (E) placed the toy into one of two opaque containers. In the more CG condition, children and caregivers jointly attended to E as E placed the toy into one of the containers. In the less CG condition, there was no joint attention (because either caregivers closed their eyes and covered their ears or left the room while E hid the toy). In the less CG condition, children more often linguistically referred to the toy or its location on the shelf compared to the more CG condition (in which case children

often just pointed). This pattern was present at the individual level, too: 11 of 15 children were more informative in the less CG condition while only 1 was less informative. These findings suggest that 2-year-olds scale their informativeness inversely with cues to CG.

In O'Neill (1996), children's linguistic reference was coded "coarsely" as having included the name of their intended toy (or not) or its location (or not). These data are insightful. Nonetheless, one might wonder: Are young children able to make finer-grained distinctions in their informativeness? Might 2-year-olds, for instance, use more pronouns when referents are jointly accessible and lexical NPs when less accessible? In Campbell et al. (2000, Study 1), 2.5- and 3.5-year-olds and E1 or E2 played games that required acting out transitive events with toys. In the more CG condition, children jointly played with E1 and the toys. In the less CG condition, children played with E2 in E1's absence. After playing, in both conditions E1 asked children questions that pulled for reference to one of the toys. (Important: E1 and children had joint visual access to the referential array of toys during E1's questioning and children's responding.) Ignoring for now the parenthetical statement, in the less CG condition, it was predicted that children would produce more informative reference to specify their intended referent, e.g., lexical NPs (i.e., because children have less CG with E1). In contrast, in the more CG condition, children were predicted to use less informative reference, e.g., pronouns (i.e., because children have more CG with E1). In contrast to these predictions, however, children were similarly informative in both conditions. What might explain this discrepancy?

Noted above in parentheses, E1 and children had joint visual access to the referential array during E1's question (Campbell et al., 2000, Study 1). Consequently, when responding to E1, children needed to (i) ignore that they were jointly attending with E1 to the array but were nonetheless merely co-present with E1 and the intended referent (see Clark & Marshall, 1981);

while (ii) recalling whether they played with E1 beforehand. Therefore, though subtle, this methodological detail required that children perform executive operations that – when combined with the task of formulating and producing linguistic reference – may have been cognitively taxing and so, potentially, equalized their informativeness between conditions.

In support of this critique, correlational data provides some evidence for a role of executive function in children's referential production. Wardlow and Heyman (2016) found that 5- to 7-year-olds' performance on a working memory task correlated positively with their tendency to appropriately use modified NPs. Relatedly, O'Neill and Miller (2013) found that 2.5- to 6-year-olds who frequently gestured while they explained their responses to E on an executive function task outperformed infrequent gesturers on that task (the task tapped set switching and inhibitory control). Relevant also are the findings by Nilsen and Graham (2009), who reported a relationship between 4- and 5-year-olds' (but not 3-year-olds') inhibitory control and their tendency to overattribute CG during referential comprehension (but not production); and Moll et al.'s (2011) report that 2-year-olds' limited inhibitory control skills may cause them to overattribute CG during referential comprehension. Taken together, these results suggest that task demands placed on children in Campbell et al. (2000, Study 1) may have masked their referring skills.

Easing the task requirements of Campbell et al. (2000), Matthews et al. (2006, Study 1) had 2.5- to 4.5-year-olds watch transitive events on a TV with E. Children were instructed to tell E1 "everything [children] could see" (*ibid*, p. 408). In the more CG condition, E1 and children jointly attended to the events onscreen while children referred. In the less CG condition, E1 was behind the TV and so lacked visual access to the events at the time of children's reference (cf. Campbell et al., 2000). Instead, children and E2 jointly attended to the videos while children referred for E1. In the less CG condition, 3- and 4-year-olds used lexical nouns more often than in

the more CG condition. Moreover, 3- and 4-year-olds often used (more informative) noun-verb combos in the less CG condition and (less informative) pronoun-verb combos in the more CG condition. In contrast, 2-year-olds were similarly informative between conditions. This suggests that 3- and 4-year-olds capitalize on interlocutor line of regard as a nonlinguistic cue to CG, while 2-year-olds do not. Interestingly, these results for 2-year-olds conflict with those of O'Neill (1996), in which 2-year-olds reliably modified their informativeness depending on cues to CG. What might explain these differing findings for 2-year-olds?

The differing findings for 2-year-olds in O'Neill (1996) and Matthews et al. (2006, Study 1) might owe, to some extent, to the different linguistic coding schemes of the two studies. O'Neill (1996) coded linguistic and nonlinguistic reference (e.g., deictic pointing gestures) whereas Matthews et al. (2006, Study 1) only coded linguistic reference. As Matthews et al. (2006, p. 421) note, this is an probably important difference, as 2-year-olds often rely on nonlinguistic referential devices to supplement linguistic reference (e.g., Tomasello et al., 1984; see Discussion Section).

Null results. It is instructive to consider reports of null effects of manipulations to listener line of regard on children's informativeness. In Bannard et al. (2017, Study 1), E1 and 3-year-olds first partook in a kind of "teaching phase" (not called that by Bannard et al.) in which children heard E1 use modified lexical NPs to refer to pictures of entities with salient properties (e.g., *pretty dress*). Then, E2 entered. E1 "prompted [children] to tell E2 about [the] pictures" (*ibid*, p. 957) as E1 showed the pictures again. In one condition, E2 and children jointly attended to the pictures while children referred to them. In another condition, E2 sat across the room, unable to see the pictures. Contrary to predictions, Bannard et al. (2017, Study 1) found no difference between conditions in children's rates of adjective use (collapsing across conditions, 79% of lexical NPs were modified). Why did children use adjectives at similarly high rates between conditions? One

possibility is that, during the teaching phase, children learned a rule that a specific construction (i.e., a modified lexical NP) is used to refer to E1's pictures. This is plausible given that children only heard examples of modified lexical NPs during teaching. Subsequently, when describing the same pictures to E2, children in both conditions may have used this rule to guide their reference.

Another set of studies was conducted by Grigoroglou and Papafragou (2019b). In this study, 4- and 5-year-olds' referring expressions included disambiguating information at similar rates regardless of whether they jointly attended to intended referents with listeners. What might explain this pattern? In test trials, the only feature that differentiated referents was the location of an agent who performed a transitive action (e.g., near a tree or a bench) or the tool the agent used to perform the action (being either a conventional or unconventional tool for the action; see below). One potential issue is that, regardless of whether interlocutors jointly attended to the referential array, both events in the array had the same actor performing the same action. Thus, to distinguish events, children had to refer to a tool or a location. That is, even though the locative (e.g., *by the bench*) and manner phrases (e.g., *with a broom*) were not syntactically obligatory adjuncts of the main verb (Grigoroglou & Papafragou, 2019b), they were discourse-pragmatically obligatory.

Conclusion. We have reviewed evidence that suggests that linguistic reference depends on listener line of regard possibly by around 2.5 years of age, likely by 3.5 years of age, and definitely by 4 to 5 years of age (4-year-olds: Matthews et al., 2006; Nilsen & Graham, 2009; 5-year-olds: Bahtiyar & Küntay, 2009; Nadig & Sedivy, 2002). The presence of nonlinguistic cues to CG causes children to be less informative while their absence causes them to be more informative. This pattern aligns with predictions of the positive view and is unaccounted for on the deficits view.

Array Configuration

All else equal, joint attention to intended referents is less likely when potential referents are more numerous, perceptually similar, or closely spaced together compared to when they are less numerous, perceptually dissimilar, or spaced apart (Clark & Marshall, 1981; Siposova & Carpenter, 2019). Consequently, speakers should infer a relatively low probability of CG about intended referents in the former compared to the latter cases. Indeed this is reflected in the finding that, compared to the latter, the former generally require more informative reference to pick out intended referents (Olson, 1970; e.g., Ford & Olson, 1975). For example, if we are seated across a table with a single candle on it, I can unambiguously refer to it simply by saying "That candle looks pretty" because to what else could I be referring (Schiffer, 1972)? And what about when there are four candles, one of which is dissimilar from the others? Still, I can say the above and you will likely identify my intended candle (Clark et al., 1983). But, when there are four identical candles or when there are two that are directly adjacent, I will need to be more informative, e.g., by saying "That candle on the left looks pretty." When do children do this?

General discussion. ***Cardinality and perceptual similarity***. The dependence of children's informativeness on referent cardinality and perceptual similarity was investigated by Matthews et al. (2012, Study 1). Two- and 4-year-olds requested stickers from arrays by asking E1 to retrieve them. E1 did not know which sticker was needed and children and E1 jointly attended to the array while children referred. In the more CG condition, children requested stickers from two-sticker arrays. In this condition, less informative reference sufficed (e.g., unmodified NPs) as the stickers were dissimilar (e.g., a sheep and a person). In the less CG condition, children requested stickers from a four-sticker array. In this condition, more informativeness was needed (e.g., modified NPs) because unambiguous identification of intended stickers required distinguishing between the intended sticker, a dissimilar sticker, and two similar stickers. Note that the proximity of the

stickers in the array relative to children's distance from the array rendered pointing ineffective. Contrary to predictions (Matthews et al., 2012, Study 1), children's reference was similarly informative between conditions. However, there were 36 trials, and E1 provided feedback about children's ambiguity when they were ambiguous. Thus, it is possible that children displayed the predicted pattern in later trials (i.e., learned from feedback). Indeed, in the final 10 trials, participants were more informative when referring to referents in arrays in the less CG condition than in the more CG condition. This suggests that, with training via feedback (see below), 2- and 4-year-olds' informativeness comes to depend on the cardinality or perceptual similarity of referents in arrays. The "or" is because the simultaneous manipulation of cardinality and similarity precludes inference about which of these cue(s) children used to scale their informativeness.

To untangle this, Matthews et al. (2012, Study 2) first gave children training to refer to stickers in four-sticker arrays. In the less CG condition, arrays contained similar stickers and so required greater informativeness. In contrast, arrays in the more CG condition had dissimilar stickers and so pulled for less informativeness. During training trials, E2 provided feedback (see below) about children's ambiguity if their reference was ambiguous. Though neither age group adapted their informativeness to the arrays during training, the generalization trials that followed gave children another opportunity to demonstrate their skills. Generalization trials tested children's ability to generalize their training with four-sticker arrays to similar or dissimilar two- or six-sticker arrays. In generalization trials, 2-year-olds' reference did not depend on referent cardinality or similarity. In contrast, 4-year-olds were more informative when referents were perceptually similar compared to dissimilar and when arrays included six compared to two referents. Overall, the results of Matthews et al. (2012) suggest that 2- and 4-year-olds calibrate their linguistic referential informativeness according to referent cardinality (also, Serratrice, 2008, Study 2) or

perceptual similarity. While it is not so clear which of these cue(s) 2-year-olds use, 4-year-olds can leverage either type of cue to scale their informativeness.

Why were 2-year-olds unable inversely scale their informativeness in Matthews et al. (2012, Study 2)? There are at least two, non-mutually exclusive explanations. One is that the perceptual dissimilarity of the referents in the less CG condition was too subtle for 2-year-olds to pick up on and thus increased listener informational needs too "opaque" (*ibid*., p. 203); perhaps referents must be very obviously distinct for 2-year-olds to leverage this cue. Second, the four-sticker training arrays may have presented 2-year-olds with too difficult a cognitive task of identifying and comparing stickers prior to reference; perhaps there must be a very small number of referents for 2-year-olds to leverage this cue. The upshot of either explanation is that 2-year-olds' referring abilities may have been masked by task demands in Matthews et al. (2012, Study 2).

Choosing one (or both) of the above explanations requires examining 2-year-olds' reference in a design that fully crosses manipulations to referent cardinality and perceptual similarity. Unfortunately, though its results are still informative, the only study to follow up on the findings of Matthews et al. (2012) did not fully cross these factors, but instead only investigated effects of referent perceptual similarity. Abbot-Smith et al. (2016) examined 2.5-year-olds' ability to adapt their reference to simplified, two-referent arrays following training via feedback about referential ambiguity. Like Matthews et al., children were exposed to more CG and less CG conditions. However, compared to Matthews et al. (2012), in Abbot-Smith et al. (2016) the number of referents that comprised the training and test arrays was smaller (two referents instead of four) and the referents were more perceptually dissimilar to one another (in the less CG condition). With these simplifications, 2-year-olds inversely scaled their informativeness with cues to CG. For

example, when presented with two identical pigs, children were more likely to supplement their NPs with additional information (e.g., *the pig on the chair*) than when presented with two different animals (e.g., *the pig*). These results suggest that 2.5-year-olds' referential informativeness depends on the perceptual similarity of referents in simple arrays (Abbot-Smith et al., 2016). Overall, we have seen that 2- and 4-year-olds inversely scale their informativeness according to referent perceptual similarity and that 4-year-olds inversely scale their informativeness according to referent cardinality. Do children also scale their informativeness according to the spatial configuration of referential arrays?

Spatial configuration. To investigate children's ability to use spatial configuration cues, O'Neill and Topolovec (2001, Study 1) gave children (2;8) a similar task as in O'Neill (1996), in which children requested that caregivers retrieve stickers hidden in one of two boxes. In the less CG conditions of O'Neill and Topolovec (2001, Study 1), linguistic reference was necessary, as the boxes were too closely adjacent to one another for unambiguous pointing. In contrast, nonlinguistic reference, e.g., pointing sufficed in the more CG conditions in which the boxes were separated. To facilitate linguistic reference, boxes were clad with images of common vehicles. As predicted, children were more informative in the less CG than more CG conditions when they requested stickers (by more often referring to a vehicle depicted on the outside of the box). Moreover, while rates of pointing were similar between conditions, participants more often produced (more informative) point + name combos in the less CG conditions and (less informative) point + deictic demonstrative (e.g., *that*) combos in the more CG conditions. While a modified version of this procedure that used the same age children resulted in a failure to replicate the findings of Study 1 (O'Neill & Topolovec, 2001, Study 2), this masked starker individual level effects. In Study 2, 56% ($N = 9$) of children more often named the referent in the less CG conditions than in the more

CG conditions, while only 2 children (12.5%) did so more often in the more CG conditions. In comparison, in Study 1 – in which group-level differences in informativeness were found – only 44% ($N = 7$) of children more often named the referent in the less CG conditions than in the more CG conditions, while 1 (6%) did so more often in the more CG conditions. Taken together, these two studies suggest that, by 2;8, children's informativeness depends on the spatial arrangement of referents. A failure to find effects in a simplified version of the above procedure with younger children (2;4) (O'Neill & Topolovec, 2001, Study 3) suggests that the months between 2;4 and 2;8 are important in the development the ability to use spatial configuration as a cue to CG.

Null results. As with discussion of listener line of regard, it is instructive to consider explanations for null results following manipulation of array configuration cues. In Wittek and Tomasello (2005, Study 3), 2.5- and 3.5-year-olds had to request one object out of three from E2. In a most CG condition, objects were widely spaced so that pronominal or null reference (plus point) was adequate. In a more CG condition, the objects were closely spaced, and thus the referent's name was required. In a least CG condition, the objects were in an opaque container – again, the referent's name was required. However, contra predictions, there were no differences between conditions in children's informativeness. This may seem surprising and counter to that found in studies reviewed above. One explanation, suggested by Wittek and Tomasello (2005), is that children, in fact, lacked CG with E2 with respect to intended referents in all conditions. That is, they never played with it together nor talked about it nor jointly attended to it. The referent was thus absent from CG and, thus, children had to introduce the referent into CG in all conditions. This explanation predicts that children should have tended to use relatively informative forms at similar rates in all conditions (i.e., to introduce the referent into CG with E2). And indeed, children often used lexical nouns, with 2.5-year-olds using them 73-78% of the time across conditions and

3.5-year-olds 90-100%. Thus, a closer inspection of the null results of Wittek and Tomasello (2005, Study 3) finds support for the predictions of the positive view. Specifically, children often use relatively informative forms when lacking cues to CG.

Conclusion. This subsection reviewed studies on the dependence of children's referential informativeness on one type of nonlinguistic cue to CG, namely, referential array configuration. Evidence was reviewed that suggested that, by around 2.5 years of age, informativeness depends on the configuration of arrays. Specifically, young children's informativeness depends on inferences about CG formed following observation of cues to referent cardinality, perceptual similarity, and spatial arrangement. This pattern is expected under the positive view but cannot be easily explained under the deficits view.

Nonlinguistic Cues to Cultural Common Ground

So far, we have discussed reference and cues to *personal common ground* (Clark, 1996; Tomasello, 2014). Inferences about personal CG require joint attention. In contrast, inferences about *cultural common ground* do not require joint attention (Clark & Marshall, 1981). Rather, cultural CG is inferred from cues to group affiliation – such as clothing or behavior – and cues to the culturally shared status of referents, such as its appearance as a conventional artifact or others' apparent familiarity with or knowledge of it (Schmidt et al., 2011, 2016). Thus, establishing that children leverage cues to cultural CG requires manipulations that target (i) interlocutor group affiliation and (ii) the cultural status of referents (e.g., Schmidt et al., 2012). Prior research suggests that 3 years of age marks a turning point in children's ability to use cultural CG for noncommunicative tasks (Tomasello, 2019). What about for their referential informativeness?

General discussion. The positive view predicts that cues to cultural CG possess a similar relationship with informativeness as do cues to personal CG. That is, the presence of more cues to cultural CG is predicted to be associated with less informative reference while fewer cues is associated with more (e.g., Schegloff, 1972; for a related point, see Kay, 1977). Unfortunately, this is only speculation: Studies that investigated the influence of cues to cultural CG on children's informativeness have thus far only manipulated cues to the conventionality of objects and not interlocutor group status (for comprehension, see Liebal et al., 2013). Moreover, sampled ages have not yet bridged the theoretically relevant age of 3 years. Thus, current evidence limits us with respect to the inferences we can draw about children's abilities to use cues to cultural CG to scale their informativeness. Nonetheless, the results of existing research are informative.

Grigoroglou and Papafragou (2019a, 2019b) asked whether the conventionality of objects used by agents to perform transitive actions influenced 4- and 5-year-olds' informativeness. The authors reported that children referred to unconventional objects more often than to conventional objects (e.g., branches rather than brooms used to sweep). This pattern is reminiscent of Greenfield and Smith's (1976) principle of informativeness, in which reference is predicted to favor encoding "informative" (unexpected) situational elements (but see Pea, 1979). In this sense, children's reference to branches is informative in that, presumably, children inferred that listeners would otherwise be licensed by cultural CG to infer that agents used brooms to sweep. Relatedly, Köymen, Rosenbaum, et al. (2014) found that 3- and 5-year-olds more often stated their reasons when justifying choices to peers about where to place unconventional compared to conventional objects in a toy zoo. For instance, when jointly reasoning about where to place a polar bear, children would say "It goes on the ice!" In contrast, when jointly reasoning about where to place a piano, children would say something more informative, like "It goes by the bench because…."

Thus, for items conventionally in zoos, children relied on cultural CG about where such items are typically located and, so, they tended not to state their reasoning. In contrast, children more often stated their reasoning when choosing where to place unconventional objects, perhaps because they recognized the lack of cultural CG (see Köymen & Tomasello, 2020).

Conclusion. There is currently limited evidence that speaks to whether young children inversely scale their informativeness according to nonlinguistic cues to cultural CG.

Interim Summary 1

This section established that nonlinguistic cues to personal CG influence the informativeness of linguistic reference beginning at around 2.5 years of age. Studies of effects of listener line of regard provide mixed evidence for the dependence of 2-year-olds' informativeness on this cue, while there is clear evidence for 3-year-olds' ability to use this cue. Arguably stronger evidence for effects of nonlinguistic cues to CG on 2.5-year-olds' informativeness came from studies that manipulated array configuration. There is limited research on children's informativeness and nonlinguistic cues to cultural CG. Overall, the findings reviewed thus far are better accounted for by the positive view, which predicts that young children's informativeness is inversely related to cues to CG. In contrast, the deficits view, which suggests that individuals are referentially incompetent until the teenage years, cannot account for the reviewed findings.

Linguistic Cues to Common Ground and Young Children's Referential Informativeness

The previous section discussed nonlinguistic cues to CG. This section turns to linguistic cues to CG. Subsection structure, condition labels, and behavioral predictions mirror those of the preceding section. The subsection on feedback poses a different set of predictions (see below).

Anaphoric Discourse

A typical referring situation for young children is as a part of stereotyped joint play with adults (Bruner, 1983). In these situations, individuals use *communicative constructions* to refer to things. Communicative constructions are statistically reliable, conventional associations between communicative form and intended meaning (Goldberg, 1995, 2003). This means that, for children, a reliable cue to CG is adults' talk – children can exploit the reliability of the constructional mapping between form and meaning to offload the cognitive work of inferring CG to the social environment (Vasil et al., 2020). Of course, exploiting constructional form as a linguistic cue to CG presupposes having learned the reliability of the mapping between form and meaning. However, as we shall see, children at least as young as 2;0 exploit the reliability of this type of cue to CG to scale their informativeness. These linguistic cues are called "anaphoric," roughly meaning "having appeared earlier in discourse." Research has investigated the effects of three types of anaphoric cues to CG, including effects of question type (e.g., Do children overtly realize NPs when presupposed by listeners' questions?), contrastive reference (e.g., Which forms do children use to refer when there are multiple, semantically related discourse referents?), and the forms listeners used (e.g., Do children use *dog* if listeners used *dog*?).

General discussion. ***Question type***. One type of anaphoric cue to CG is the type of question interlocutors use to prompt children's reference. Generally, investigations of the effects of question type on informativeness focus on manipulations to question information structure (Lambrecht, 1994). A typical manipulation is one in which children are asked questions that suggest that their interlocutor shares either more CG (predicate-focus questions) or less CG (sentence-focus questions). Predicate-focus questions are used to inquire into predicate argument(s) and often

include cues to the CG status of inquired entities (e.g., definite reference; see Schmerse et al., 2015). For example, the predicate-focus question of the declarative sentence *the dax pushed the ball* would be *what did the dax do?* Felicitous response to this type of question does not require lexical reference to subject NPs. That is, one only needs to use null or pronominal reference to refer to the subject when responding to predicate-focus questions, e.g., *pushed the ball* or *it pushed the ball*. In contrast, sentence-focus questions suggest less CG about intended referents. A typical sentence-focus question is *What happened?* Felicitous response to this type of question requires greater informativeness compared to responses to predicate-focus questions. In particular, responses to sentence-focus questions typically require reference to subjects using lexical NPs or pronouns, e.g., *the dax* in *the dax pushed the ball*. When do children use question type as a cue to CG?

That was the research question of Campbell et al. (2000, Study 2). These authors investigated whether 2.5- and 3.5-year-olds scaled their informativeness according to the type of question E asked them. Children and E played with a toy before E asked children about the toy. In the more CG condition, E asked children about the toy using predicate-focus questions. In the less CG condition, E asked about the toy using sentence-focus questions. Note that, because children and E jointly attended to the toys beforehand (i.e., during play), predicate-focus questions pulled for null reference (e.g., *fell*). In contrast, sentence-focus questions pulled for children to use lexical NPs or pronouns (e.g., *the ball fell*, *it fell*). When responding to E's questions, children of both age groups followed the predicted pattern. Specifically, children provided more informative reference (lexical NPs and pronouns) in response to sentence-focus compared to predicate-focus questions (in which case they favored null reference). These findings align with the predictions of the positive view. In particular, these findings suggest that even 2.5-year-olds inversely scale their

informativeness for listeners according to cues to CG present in anaphoric discourse, namely, question type.

In Campbell et al. (2000, Study 2), joint attention to intended referents rendered them recoverable from nonlinguistic context. Consequently, in that study, pronominal reference was felicitous in response to sentence-focus questions. But what about when referents are not fully recoverable from nonlinguistic context? It may then be that pronouns are suitable only in response to predicate-focus questions while lexical NPs and nulls retain their felicity profiles. One might therefore ask about interactions, namely, whether children's tendency to use pronouns in response to different types of question depends on nonlinguistic cues to CG. One way to investigate this is indirectly, that is, in experiments with comparable designs and manipulations to cues to CG as in a baseline study but in which there is an (unmanipulated) change in the level of the other class of cues to CG.

In this sense, Wittek and Tomasello, (2005, Study 1) indirectly investigated whether children's use of pronouns in response to different question types depends on nonlinguistic cues to CG. Children and E jointly attended to an array of toys while E asked children which toy was needed to play a game. Importantly, in contrast to Campbell et al. (2000, Study 2), interlocutors in Wittek and Tomasello (2005, Study 1) had in their CG that E did not know which toy was the intended referent (i.e., because there was only joint attention to the referential array, not to referents within the array). In the less CG condition, when E asked about the toy (*what do we need to get?*), 2.5- and 3.5-year-olds often responded with lexical nouns (but not 2.0-year-olds; see Wittek & Tomasello, 2005, Study 2). In contrast, in the more CG condition, when E asked about the toy (*what happened to the broom?*), children often used pronominal and null reference. Key here is the difference in children's use of pronouns in the more CG and less CG conditions between Wittek

and Tomasello (2005, Study 1) and Campbell et al. (2000, Study 2). In the former study, children understood that pronominal reference was infelicitous in response to sentence-focus questions (i.e., because there was no joint attention to intended referents). Meanwhile, in the latter study, children understood that pronominal reference was felicitous in response to the same type of question (i.e., because there was joint attention to intended referents). Taken together, the results of this pair of studies may suggest that 2.5- and 3.5-year-olds understand that the informativeness pulled for by linguistic cues to CG depends on nonlinguistic cues to CG.

So far, the focus has been on indirect investigations of interactions between classes of cue to CG. This choice of focus is due to difficulties in interpreting the results of existing direct investigations of such dependencies (i.e., research that has crossed the presence/absence of linguistic and nonlinguistic cues within one design). Of the two direct investigations of question type and nonlinguistic cues to CG, one indicated main effects of question type and listener line of regard on informativeness but no interaction (Serratrice, 2008, Study 1), while one found main effects of listener line of regard but not for question type nor an interaction (Salomo et al., 2011). While these studies might suggest that nonlinguistic and linguistic cues to CG operate independently, this suggestion can only be tentative. In the within-subjects design of Serratrice (2008, Study 1), it is unclear whether the order of presentation of levels of the linguistic cues manipulation was counterbalanced, and no analysis of order effects for this manipulation is reported; it is thus possible that the effects of linguistic cues were spurious, due to condition order. Salomo et al. (2011) remedied that issue but reported that children were insensitive to question type until age 4. This contrasts with the results of Serratrice (2008, Study 1), who reported that 3-year-olds were sensitive to question type, and with Campbell et al. (2000, Study 2) and Wittek and Tomasello (2005, Study 1), who reported the same for 2.5-year-olds. One explanation is that

children in different studies referred to entities with different semantic and syntactic roles in different argument structure constructions (Salomo et al., 2013). Children in Salomo et al. (2011) referred to transitive patients, those in Serratrice (2008, Study 1) to transitive agents, and those in Campbell et al. (2000, Study 2) and Wittek and Tomasello (2005, Study 1) to intransitive agents and locations, respectively. Overall, existing research paints an unclear picture of interactions between question type and nonlinguistic cues to CG on children's informativeness.

Recall that, in the procedure of Wittek and Tomasello (2005, Study 1), interlocutors jointly attended to the referential array (but not intended referents) when E asked children the target question. However, in the previous section, it was noted how joint attention to referential arrays is not the same as intended referents being out of sight. This is because referents are still co-present with interlocutors as part of referential arrays. Therefore, it is unclear from the results of Wittek and Tomasello (2005, Study 1) whether children are similarly skilled at calibrating their informativeness in settings that eliminate nonlinguistic cues to CG, that is, in which there is neither joint attention to intended referents nor to referential arrays. This is important to ask because very young children's ability to form inferences about CG (and, by extension, their ability to scale their informativeness) may benefit from the presence of more cues to CG. Do children rely on question type to scale their informativeness when nonlinguistic cues to CG are eliminated?

This question was addressed by Matthews et al. (2006, Study 2). In a key methodological departure from Campbell et al. (2000) and Wittek and Tomasello (2005), the referential array in Matthews et al. (2006, Study 2) was out of sight of E when E asked their questions and children responded. Thus, there was joint attention to neither intended referents nor arrays. Two- to 4-year-olds described the contents of videos to E while children watched them; E was behind the TV, unable to see the videos. In a less CG ('sentence-focus') condition, E said "That sounds like a fun

video. What happened?" Meanwhile, in a more CG ('predicate-focus') condition, E said "Was that the clown? What happened?" Thus, while both questions, technically, were sentence focus, the immediately preceding discourse was manipulated so that E, in effect, asked about the video or the clown in the video (hence the '', above). When responding in the more CG condition, 3.5- and 4.5-year-olds used pronouns more often than in the less CG condition. Thus, the 3-and 4-year-olds in Matthews et al. (2006, Study 2) responded like the 3-year-olds in Wittek and Tomasello (2005, Study 1). In contrast, 2.5-year-olds in Matthews et al. (2006, Study 2) used lexical nouns, pronouns, and null reference at similar rates in both CG conditions. Thus, when neither referential arrays nor intended referents were jointly attended, 2-year-olds were unimpacted by question type. These results contrast with 2.5-year-olds' behavior in situations in which referential arrays (Campbell et al., 2000, Study 2) or intended referents were jointly attended (Wittek & Tomasello, 2005, Study 1). What might explain this?

One explanation falls on the difficulty children have in reconciling conflicting perspectives (Tomasello, 2018). Only in Matthews et al. (2006, Study 2) did children have to fully inhibit their individual (visual) perspective on the referent and take into account the shared perspective invited by anaphoric discourse. In contrast, the cues to which children were exposed in Campbell et al. (2000, Study 2) and Wittek and Tomasello (2005, Study 1) did invite fully conflicting perspectives (because of joint attention to intended referents or to referential arrays, respectively). This is probably an important methodological detail, as it is not until 3 to 4 years of age that children can simultaneously compare conflicting perspectives and act on the outcome of that comparison (Moll et al., 2013). Thus, a difficulty in reconciling conflicting perspectives may have played a role in the difficulty 2-year-olds had in distinguishing between conditions in Matthews et al. (2006, Study 2) compared to Campbell et al. (2000, Study 2) and Wittek and Tomasello (2005, Study 1).

Contrast. Interlocutors often converse about multiple, semantically similar referents at once. This presents challenges for unambiguous reference because intended referents must be distinguished from other referents that are also highly accessible. Adult speakers possess various strategies to preempt such ambiguity, e.g., using more informative arguments to distinguish intended from unintended referents (Givón, 1992; e.g., Karimi, 2022). How do young children navigate this challenge?

Above, we discussed 2.5- and 3.5-year-olds' responses to less CG (sentence focus) and more CG (predicate focus) questions in Wittek and Tomasello (2005, Study 1). A third level of this manipulation was also present. In a contrast question condition, E asked whether an incorrect toy was required to play a game with participants (e.g., *Did the clown have a vacuum cleaner?*). The idea was that, because there were multiple referents in the jointly attended referential array (i.e., other toys), unambiguous reference required children to respond by using lexical NPs to distinguish their intended toy from the others (e.g., *No, a/the broom*). Thus, children's responses in the contrast question condition were expected to look more similar to their responses in the less CG condition than in the more CG condition. And indeed, when responding to E, children in the contrast question condition used lexical nouns at a similar rate as in the less CG condition and at a higher rate than in the more CG condition. Moreover, whereas null reference was the predominant response in the more CG condition, children never used null reference in the contrast condition. These results suggest that anaphoric discourse that includes multiple referents is a linguistic cue to CG that 2.5-year-olds use to scale their informativeness.

Indeed, similar results have been reported for even younger children. In Salomo et al. (2010, Study 1), E narrated for children (2;3) as both jointly watched an agent either perform multiple actions on a patient (*patient CG condition*) or repeatedly perform an action on multiple

patients (*action CG condition*). E then asked children questions that prompted them to refer to different aspects of the situation. In the patient CG condition, children had to distinguish one among multiple actions; this required referring to the action alongside pronominal or null reference to the patient. In the action CG condition, children had to distinguish one among multiple patients; this required referring to the patient with a lexical NP. Children demonstrated their understanding of this by referring to actions more often in the patient CG than action CG condition and to patients more often in the action CG than patient CG condition. Moreover, when they referred to patients, children tended to use null reference in the patient CG condition and lexical NPs in the action CG condition. Indeed, still younger children have been found to scale their informativeness for contrastive purposes. Wittek and Tomasello (2005, Study 2) found that children aged 2;0 more often used lexical NPs following contrast compared to predicate focus questions (in which case they tended to use null reference). Altogether, these results suggest that children as young as 2;0 inversely scale their informativeness during contrastive reference.

A running thread in the present section has been the interplay of nonlinguistic and linguistic cues to CG. So far, important takeaways are that there is currently limited evidence for interactive effects between the two classes of cues to CG; and that 2-year-olds' ability to scale their informativeness may suffer when conflicting perspectives are simultaneously invited by either class of cues to CG. Now, recall that, in Campbell et al. (2000, Study 2), Wittek and Tomasello (2005, Study 1), and Salomo et al. (2010, Study 1), children and E always supplemented their discourse about intended referents with joint attention to them or to referential arrays. Following the thread being woven in this section, one might ask about the role of nonlinguistic cues to CG in these studies.

To investigate this, Salomo et al. (2010) conducted a second study. As in Salomo et al. (2010, Study 1), children (2;3) jointly attended to transitive events with E before E questioned them. However, this time, children did not hear contentful anaphoric discourse (instead, E said things like "That's great."). This ended up being an important methodological change. While children still referred to the action in the patient CG condition more often than in the action CG condition (like Study 1), they no longer referred more often to the patient in the action CG condition than in the patient CG condition (unlike Study 1). Moreover, while use of lexical NPs to refer to patients was more prevalent in the action CG than in the patient CG condition (like Study 1), lexical NPs were used less often in the action CG condition of Study 2 than Study 1. These results do not license inferences about whether children preferably rely on linguistic over nonlinguistic cues to CG. However, these results do suggest that 2-year-olds are less able to modify their informativeness when only nonlinguistic cues are present compared to when nonlinguistic and linguistic cues are present.

Taken together, the results reviewed above suggest that (i) children rely on linguistic before nonlinguistic cues to CG and (ii) modify their informativeness at especially young ages for contrastive reference. Indeed, children's ability to calibrate their informativeness for contrastive purposes quickly becomes rather sophisticated: By the time they turn 3, children omit superfluous subjects, verbs, or objects depending on whether they are contrasting semantically compatible agents, actions, or patients (Graf et al., 2015). This pattern is accounted for under the positive view of early referential production and is unaccounted for under the deficits view.

Null results. Informative null results come from studies that investigated effects of anaphoric referring forms on informativeness. Campbell et al. (2000, Study 1) asked whether an E's reference to a toy using a common noun only, or a common noun and pronoun, had any effect on 2.5- and

3.5-year-olds' informativeness ("an E" because whether E1 or E2 referred was manipulated, too; see above). After playing with the toy with an E, E1 questioned children about the toy. In their responses, children used lexical NPs, pronouns, and null reference at similar rates regardless of the anaphoric forms they heard an E use. While this null result might seem surprising (and was counter the predictions of the authors), on the positive view this pattern of behavior is unsurprising. This is because there is no reason to think that the use by an E of common nouns only, or common nouns and pronouns, actually altered the CG status of the referent itself for participants and E1 (and, for this reason, we have ignored the "more CG" or "less CG" naming convention for this manipulation in this study). Rather, the CG status of the referring form was manipulated, that is, the "referential pact" shared between interlocutors about how to refer to CG entities (Brennan & Clark, 1996). The issue is that children's informativeness does not depend on referential pacts until around 6 years of age (Köymen, Schmerse, et al., 2014). Therefore, the 2- and 3-year-olds in Campbell et al. (2000, Study 1) may not have been impacted by the manipulation because it targeted a phenomenon that does not characterize the behavior of children that young.

Campbell et al. (2000, Study 2) asked whether the cost of production of referring forms influences children's referential choices. The authors investigated the informativeness of 2.5- and 3.5-year-olds' responses to E's questions about a toy's action that children caused (e.g., a ball falling). In anaphoric discourse, E referred to toys with one of four types of nominal: Simple familiar (e.g., *ball*), simple novel (e.g., *toam*), complex familiar (e.g., *elephant*), or complex novel nouns (e.g., *figertruckle*). While it was predicted that children would tend to produce (relatively less effortful) non-lexical referring forms when referring to entities that E referred to with novel or complex forms, children were in fact similarly informative across conditions. Why was this?

There are two possible explanations. One is the same as above, namely, that the CG status of referents *per se* was not manipulated in Campbell et al. (2000, Study 2). Indeed, that children tended to use pronouns regardless of the form(s) that E used is evidence for this explanation. A second explanation involves factors not investigated in Campbell et al. (2000, Study 2). Bannard et al. (2017) found that 3-year-olds more often used modified NPs the more informative were the modifiers used by E in a "teaching phase" (see above). However, as modifiers became more informative, children became less likely to imitate them and, instead, switched to semantically similar, though less informative modifiers. Extrapolating to the use of lexical nouns and pronouns, one possibility is that children more often use lexical nouns when E refers with highly informative lexical nouns and more often use pronouns when E refers with less informative lexical nouns. Following the extrapolated pattern, the issue, then, is that E's referring forms were equally informative in all four conditions of Campbell et al. (2000, Study 2). This, in turn, may have masked effects of cost of production on children's choice of using lexical nouns or pronouns.

Conclusion. The informativeness of linguistic reference depends on linguistic cues to CG by 2;0. In contrast, it is not until 2;6 that informativeness depends on nonlinguistic cues to CG. It is unclear whether nonlinguistic and linguistic cues to CG exert interactive effects on children's informativeness. Two-year-olds may have difficulty scaling their informativeness when nonlinguistic and linguistic cues to CG invite conflicting perspectives on intended referents. These findings support the positive view and do not support the deficits view, as the latter suggests that children are referentially incompetent until the teenage years.

Listener Feedback

Reference is a collaborative process (Clark & Wilkes-Gibbs, 1986). Speakers refer and listeners provide feedback, including expressions of perceived understanding (*ok!*) or misunderstanding (*what?*) or comments on intended referents. Feedback enables speakers to infer whether they have successfully contributed to conversation by having placed intended referents into CG (Clark & Schaefer, 1987). If feedback indicates unsuccessful reference, speakers may produce repaired reference that clarifies the misunderstanding. Indeed, by the time they are 2 years old children produce repaired reference following feedback (Grosse et al., 2010). Moreover, young children repair their reference on the basis of rather uninformative feedback, like facial expressions of sadness (Bacso et al., 2021). What about more explicit cues, like listener feedback questions?

When asking this question, it is important to distinguish the cognitive demands of initial versus repaired reference. Only during repaired reference do children have the benefit of hindsight, namely, that initial reference failed and that more informativeness is required. Crucially, in this context, the effects of cues to CG on informativeness should be *opposite* the inverse pattern discussed so far. That is, the positive view predicts that feedback cues that indicate more CG should favor more informative repaired reference compared to cues that indicate less CG. This is because there is less cognitive work that children have to do to identify the source of referential failure when feedback indicates the source of the failure. This frees up cognitive resources to construct more informative repaired reference. This pattern is illustrated by children's behavior in prior research.

General discussion. Two studies have investigated children's ability to use feedback questions as cues to CG. Three- and 4-year-olds in Nilsen and Mangal (2012) received either goal substitution (i.e., E handed children an unintended sticker), explicit statement of misunderstanding (i.e., E said *I don't know what you mean*), or vague feedback following ambiguous initial reference (i.e., E said

huh?). Arguably, the goal substitution condition is like a predicate-focus question in that interlocutors' joint attention is focused on an (unintended) referent, while the latter two are more or less (resp.) verbose variants of sentence-focus questions. Additionally, in the goal substitution condition listeners displayed uptake of children's requestive speech act (even if listeners misidentified intended referents) but displayed no such uptake in the other two conditions. Consequently, the goal substitution condition is labeled the more CG condition and the other two conditions the less CG conditions. Assuming this analysis of the conditions, the pattern predicted by the positive view characterizes children's behavior. Children in the more CG condition produced more informative repaired reference (i.e., more descriptor terms) than children in the two less CG conditions (Nilsen & Mangal, 2012). Moreover, informativeness was similar in the two less CG conditions, as expected under the analysis of both as sentence-focus questions.

The results above support the positive view. However, this support is weak because the more CG condition really only implicated nonlinguistic cues to CG. Would children behave as they did in Nilsen and Mangal (2012) if only linguistic cues to CG were manipulated? Unfortunately, no other studies sampled the same age range that speak directly to this issue. However, Bacso and Nilsen (2017) found that 4- to 6-year-olds provided more informative repairs, and repeated their initial reference less often, in a more CG compared to less CG condition (more CG: *There are two boys with red shirts and I don't know which one you mean*; less CG: *I don't know which one you mean*). Interestingly, Bacso and Nilsen (2017) also found that children's executive functioning (specifically, cognitive flexibility) correlated with the informativeness of their repaired reference only in the less CG condition. Arguably, these findings support our claim that the added cognitive work involved in referential repair with fewer cues to CG limits children's ability to be informative compared to when sources of referential failure are indicated in feedback.

Additional results. On first blush, additional research seems suitable for inclusion in the general discussion. However, this research investigated something other than what was investigated by studies reviewed in the general discussion. There are two paradigms used to study feedback and reference. One paradigm is what might be called the "inference paradigm." Studies that used this paradigm were reviewed in the general discussion and examined how feedback influences children's inferences and their immediate referential repair. While children may learn from their experiences in the inference paradigm, their behavior depends not on their learning but on their online inferences about CG. In contrast, a complementary "learning paradigm" examines how children learn to use referring strategies and skills from prior experience with feedback. In the learning paradigm, children participate in training and test phases during which they repeatedly refer for E (e.g., Abbot-Smith et al., 2016; sometimes, there is only an extended test phase that, due to its length, *de facto* includes training; e.g., Matthews et al., 2012). When children's reference is ambiguous, E provides feedback that indicates more or less CG. However, the dependent variable is not children's repaired reference, but their initial reference in subsequent trials. Consequently, one is investigating a learning rather than an inference phenomenon in the learning paradigm.

This is not to say that the learning paradigm is irrelevant for us. In fact, a coherent picture of effects of prior experience with feedback on the informativeness of initial reference emerges from this work. Abbot-Smith et al. (2016) found no effect among 2.5-year-olds of predicate-focus (more CG) compared to sentence-focus (less CG) feedback training on initial informativeness. However, as above, the results of Matthews et al. (2012, Study 1) suggest that children's inability to benefit from feedback in Abbot-Smith et al. (2016) may have been due to an insufficient number of training trials. Matthews et al. (2012, Study 1) gave children 36 trials in which they could receive

feedback about their ambiguity. In the final 10 trials, but not earlier, 2- and 4-year-olds were more informative in the more CG condition than they were in the less CG condition. (Note that Abbot-Smith et al. report a strong numerical trend in this same direction, but it was not statistically significant.) In sum, although the learning paradigm investigates a different phenomenon than the inference paradigm, children's initial reference is more informative following feedback with more cues to CG than fewer cues to CG, like their repaired reference in the inference paradigm.

Conclusion. Feedback is a special kind of cue to CG compared to other cues to CG reviewed in this article. Due to limitations of children's cognitive capacities and skills, the positive view predicts that informativeness following feedback should be opposite the inverse pattern predicted for other cues. Research was reviewed that supported this prediction. By at least 3 to 4 years of age (and possibly earlier; e.g., Anselmi et al., 1986), the informativeness of repaired reference depends on feedback cues to CG. Also, 2-year-olds' initial reference benefits from prior learning experiences that include feedback that indicates the sources of referential failure.

Linguistic Cues to Cultural Common Ground

General discussion. To the present author's knowledge, there is no experimental research on linguistic cues to cultural CG and children's informativeness. There are several routes forward. Are children less informative for others with whom children infer they share cultural CG? For example, accent is an indicator of group affiliation (Cohen, 2012), and young children selectively expect to share CG with native accented speakers of children's language (Schmidt et al., 2012). Does this expectation influence informativeness? For instance, are children more informative for non-native accented compared to native accented listeners? Suggestively, Diesendruck (2005) found that 3-year-olds' interpretations of a speaker's reference depended on speaker accent.

In an influential paper, Prince (1981) proposed the notion of "inferrables." Entities or relations are inferrable if one conceptually contains another, such that reference to the containing entity increases the cognitive accessibility of the contained entity. Cultural CG roles are often inferrable. For instance, in a society with buses and bus drivers, I can felicitously use definite reference to refer to the bus driver while telling you about a bus I was on, despite only the bus – and not the bus driver – having prior mention in discourse (e.g., *A crazy thing happened on the bus. The driver uppercutted a passenger square in the face*). Does children's informativeness depend on cultural CG inferrables? This may be investigated by manipulating the linguistic cues used to teach children novel inferrables, e.g., whether anaphoric discourse includes labels used to refer to tools (contained entity) in novel games (container entity; e.g., Rakoczy et al., 2008).

Conclusion. There is no research on linguistic cues to cultural CG and children's informativeness.

Interim Summary 2

This review considered how young children inversely scale their informativeness with cues to CG (save for feedback, in which case the relationship is a positive one). The literature suggests that informativeness depends on linguistic cues to CG at a younger age than nonlinguistic cues to CG (around 2;0 and 2;6, respectively). It was suggested that executive skills needed to reconcile conflicting perspectives play a role in children's ability to scale their informativeness when cues to CG invite conflicting perspectives on referents; and that repaired informativeness benefits from feedback that indicates the source of referential failure (due to less cognitive work needed to identify the source). In all, children's referring skills are better captured by the positive view than by the deficits view, which discounts the idea that children can modify their informativeness for listeners. The following section elaborates the model that motivates the positive view.

A New Look at Young Children's Referential Informativeness

Collaboration is the ontogenetic home of communication and reference (Bruner, 1983; Clark, 1996; Levinson, 2006; Schegloff, 2006; Tomasello, 2008). It is in face-to-face collaborative activities, structured by joint goals and joint attention, that reference is first used to coordinate mental states. In fact, collaboration is a potent source of motivation for children (Tomasello, 2019), e.g., linguistically framing social activities as collaborative rather than individualistic increases children's motivation to complete those activities (Vasil, accepted; e.g., Vasil & Tomasello, 2022). Because reference is intentional behavior, it is reasonable to ask whether the motivating influence of collaboration fosters children's willingness to put in the effort to be informative.

This was investigated by Grigoroglou and Papafragou (2019b; also, 2019a). These authors compared children's informativeness in noncollaborative and collaborative contexts. Four- and 5-year-olds identified target from nontarget events for listeners. Unambiguous reference required referring to tools that agents used to perform actions or to agents' locations. In the noncollaborative context (Study 1), interlocutors were informed separately of their goals and roles in the task and listeners were introduced as E's friend. In contrast, in the collaborative context (Study 2), interlocutors were informed together of their goals and roles and listeners were introduced as children's partner. As collaborators, children were more likely than as noncollaborators to mention tools and locations (and agents, which were irrelevant for unambiguous reference). Moreover, children produced longer utterances as collaborators, with larger gains for 4- than 5-year-olds. Our and Grigoroglou and Papafragou's (2019b) conclusion is that collaboration stokes children's cooperative motivations and this, in turn, causes them to put in the effort to be more informative.

In this article, three sets of findings have been reviewed. The first set is that initial informativeness is inversely related to cues to CG; the second is that repaired informativeness is positively related to feedback cues to CG; and the third is that collaboration increases children's motivation to be informative. The first two sets inform a conclusion about a cognitive phenomenon (i.e., inferences about CG) and the latter informs a conclusion about a motivational phenomenon (i.e., the incentivizing effects of collaboration). The deficits view cannot account for the former two sets of findings and ignores motivation and, thus, cannot account for the third (Glucksberg et al., 1975). In contrast, the positive view accounts for all findings. To demonstrate this, the model of human communication that motivates the positive view is elaborated as follows.

General discussion. Vasil et al. (2020) cast human communication as a process of active inference. Active inference is a formulation of Bayesian belief-guided control in biological systems (Parr et al., 2022). There are two ways to construe active inference. One construal casts active inference as an enactive formulation of the Bayesian brain (Knill & Pouget, 2004), that is, equipped with motor reflexes (Friston, 2010). A complementary construal begins by asking how organisms individuate themselves from their environment over various timescales (Friston, 2012).

Beginning from the second construal, start with a graph (a set of nodes and edges) that is separated into *internal states* (children's neurocognitive states) and *external states* (the child-external setting). Separation invokes a statistical construct called a *Markov blanket* (Pearl, 1988; Friston, 2012). Markov blankets separate internal from external states by inducing independencies between them, when conditioned on third set of states called blanket states (Figure 2). Blanket states are defined by directed causal interactions between external and internal states. "Interaction" here can be read as the bidirectional action of internal states on external states, and external states on internal states, via blanket states. Internal states influence external states via active states of the

blanket (i.e., the neuromuscular system as used in linguistic reference). In turn, external states influence internal states via sensory states of the blanket (i.e., the sensory modalities involved in the transduction of cues to CG). Blanket states thus circularly couple internal and external states and thereby enable an exchange of information in the spirit of an action-perception cycle.

In resisting entropic dispersion, biological systems necessarily frequent a limited number of states (e.g., heart rate fluctuates within a set zone). These states are *a priori* expected, or phenotypical, states that are characteristic of the creature in question (Ramstead et al., 2018). For humans, phenotypical states include one in which conspecifics' mental states are aligned (Tomasello, 2014, 2019). Maintenance of the phenotype requires action that (preemptively or retroactively) counters unexpected fluctuations (Bruineberg et al., 2018). The Markov blanket renders this a process of (approximate Bayesian) inference, in which the dynamics of internal states entail a *generative model* that relates external states – hidden behind the Markov blanket – and observations (Friston, 2012). In active inference, internal states come to encode approximate Bayesian beliefs about how hidden states cause sensation, under a generative model. Hidden states include external states that organisms do and do not control, respectively. These states are "hidden" because they are inferred via sensation. Thus, phenotypical states are maintained by inferring the causes of sensory states, which include active states (i.e., policies; see below) and external states. The former entails acting and the latter perceiving in ways that garner evidence for the generative model. Prospective inference about optimal actions makes it *active* inference and has been referred to as planning as inference (Botvinick & Toussaint, 2012) or self-evidencing (Hohwy, 2016).

Vasil et al., (2020) propose that humans characteristically expect to encounter sensory outcomes indicative of the alignment of hidden mental states (i.e., the internal states of others), namely, cues to CG. This is cast as a species-typical *adaptive prior belief* (Badcock et al., 2019)

that mental states are aligned with others "like me," that is, that CG exists. This *adaptive prior for alignment* motivates patterned action-perception cycles that maximize evidence for the adaptive prior, on average across time. In other words, the adaptive prior for alignment energizes a uniquely human kind of self-evidencing across various timescales, ranging from dyadic interactions through to neurodevelopment and cultural niche construction (also, Veissière et al., 2020).

The goodness of fit of generative models to data is scored by their *variational free energy* (Parr et al., 2022). Lower (higher) free energies indicate that generative models are better (worse) explanations for data. There are several mathematically equivalent definitions of free energy (see Parr et al., 2022). One definition highlights the importance of perception in free energy minimization (via model inversion; Dayan et al., 1995). This definition states that models with low free energy encode simple and accurate explanations of hidden states. Increasing explanatory complexity incurs computational and metabolic cost and so is automatically penalized, though may be required for sufficiently accurate explanations (Ortega & Braun, 2013). Under the adaptive prior for alignment, free energy minimization begets relevance maximization (Sperber & Wilson, 1986), whereby "an assumption with greater contextual effects [accuracy] is more relevant; and… an assumption requiring a smaller processing effort [complexity] is more relevant," (*ibid*., p. 124).

A second definition highlights the importance of action in free energy minimization. This definition casts free energy as the sum of a divergence term (the expected difference of the approximate and exact posteriors over hidden states) and *surprisal*. Surprisal is the negative log probability of outcomes under a generative model. Heuristically, surprisal scores the unexpectedness of sensation under a generative model (Dayan et al., 1995). The important point is that the divergence term is an upper bound on surprisal. Thus, minimizing the divergence term (via perception that optimizes beliefs about hidden states) implicitly minimizes surprisal and, thus,

free energy. Functionally, minimizing the divergence term renders internal states a "good" (useful) model of external states (Conant & Ashby, 1970). Consequently, action can be leveraged to elicit expected sensations. For example, Vasil et al. (2020) argue that, across the timescale of interaction, cues to CG drive belief updates that cause speakers to converge on a shared generative model, that is, a model of the "true" CG. With convergence, interlocutors are better able to predict each other and thereby minimize the (joint) surprisal of "our" sensory exchange (Friston & Frith, 2015). Consequently, speakers can adaptively engage with the shared process (e.g., refer felicitously) because they embody a model of that process, that is, of what "we" share in CG.

The findings reviewed in this article suggest that children's reference relies on generative models that support hierarchically and temporally deep beliefs (Figure 3). In hierarchical models, the activity of internal states (e.g., neural representations) in higher levels conveys beliefs to lower levels (Friston, Rosch, et al., 2017; e.g., Friston et al., 2020). These beliefs are, in effect, allostatic and homeostatic set points that guide probabilistic transitions among hidden states at lower levels, including states that map to action (Pezzulo et al., 2015). This leads action to look as if it were gathering evidence for prior beliefs, on average across time (Friston, FitzGerald, et al., 2017).

Under the adaptive prior for alignment, beliefs that CG exists furnish prior constraints over lower hierarchical levels. These prior beliefs guide action-perception cycles to gather evidence for the existence of CG. One such action is reference. That is, reference emerges as an embodied tool to gather evidence for the adaptive prior for alignment that both depends on – and influences – beliefs about CG. Beliefs about CG probabilistically bias competition among the unobservable half of communicative constructions (Cisek, 2007), a set of states called *conventional symbolic unit* (CSU) states (Langacker, 1987). CSU states probabilistically map to phonetic output, i.e., the observable half of constructions. While CSU states may themselves possess hierarchical structure,

key is that CSU states are often linked to other CSU states at the same level of the generative model (Bybee, 2002; Frank et al., 2012). These latter are an important kind of *communicative policy* that generate extended sequences of patterned communicative action. Communicative policies may be used with referential intent, in which case they are *referential control policies*.

When planning what to do next, plausible policies are evaluated in terms of their *expected free energy*. This is the free energy that would be expected when selecting a particular plan or course of action. The expected free energy of a policy is inversely proportional to its log probability (Friston, FitzGerald, et al., 2017). More plausible policies have lower expected free energies. Whereas variational free energy is based on present and past sensation, expected free energy is based on expected sensations under competing policies. As with variational free energy, there are several mathematically equivalent decompositions of expected free energy (Parr et al., 2022). One decomposition shows that expected free energy is a mixture of *epistemic*, *pragmatic*, and *deontic value* terms. An extended example illustrates the meaning of and relationship between these terms.

In referential communication tasks, children may first infer whether they are in a collaborative or noncollaborative context. Such actions possess *epistemic affordance* by virtue of resolving contextual uncertainty, e.g., via saccades that maximize expected information gain (Poli et al., 2020). Under the adaptive prior for alignment, an important effect on behavior of inferring that one is in a collaborative context is motivational (Pezzulo et al., 2018). Collaboration motivates epistemic behavior that searches further for contextually relevant cues to CG that resolve uncertainty about CG (e.g., Mirza et al., 2019) before exploiting updated beliefs about CG to secure preferred sensory exchanges (i.e., cues to CG that indicate that mental states are aligned due to referential control). Discussed below, in this setting cues to CG come to acquire deontic value; namely, the (learned) likelihood of policies given particular sensory cues (Veissière et al., 2020).

Collaboration is hypothesized to have this motivating effect because collaborators are *a priori* more likely to possess CG than are noncollaborators. Minimally, to conceive of themselves as such, collaborators must infer that it is in their CG that they are collaborating, whereas the opposite makes no sense for noncollaborators (i.e., because CG can only be formed via communication, including nonlinguistic but ostensive looks, and human communication is always collaboration; Clark, 1996; Tomasello, 2008). As soon as individuals learn this contingency (probably via recurrent joint engagement in infancy), subsequent collaboration licenses the inference of something like "I am on the right track (i.e., towards alignment), given my adaptive prior." This inference motivates subsequent control because – given this inference – the expected benefits outweigh the expected costs of control. In other words, collaboration is an incentivizing cue that motivates intentional control to access preferred states (Berridge, 1996). Thus, the present model predicts, first, that young children should be motivated to collaborate (e.g., Rekers et al., 2011) due to the learned (deontic) value of collaboration. Second, the model predicts that, within the inferred collaborative context, children should be more motivated to engage in effortful control that (i) seeks out cues that resolve uncertainty about CG and (ii) generates more informative reference. Claim (ii) accounts for the first set of findings, in which children were more informative in collaborative than in noncollaborative contexts. What about other two sets of findings?

To answer this, first note that attention to cues to CG boosts their influence on belief updates by causing them to ascend to higher levels the generative model (Ramstead et al., 2016). Updated beliefs about CG calibrate the probability distribution over competing referential control policies. Sometimes, policies with greater deontic value will be more probable (Constant et al., 2021; Veissière et al., 2020). At other times, policies with greater pragmatic value will be more probable. Deontically valuable policies include cheap, automatized referential control policies that

require relatively little constructive effort to plan and produce. In contrast, pragmatically valuable policies usually require greater constructive effort to plan and produce and tend to be less practiced, though may be useful for shoring up certainty that one will elicit preferred sensations. What shapes the probability that pragmatic or deontic referential control policies guide action?

Deontic and pragmatic value share at least two kinds of relationships. The first speaks to the second set of findings (i.e., that cues to CG possess an inverse relationship with initial informativeness). If cues to CG increase children's confidence about the existence of CG, they might favor deontically valuable policies (e.g., nulls or pronouns). Put differently, observing cues to CG might license the automatic inference of something like "I see cues to CG → use a pronoun." Thus, cues to CG are a kind of *deontic cue* that pull for the use of relatively cheap forms by way of reliably indicating CG (Constant et al., 2019). In contrast, when children are relatively less confident about CG (such as when there are fewer or relatively unreliable cues to CG), they might favor pragmatically valuable policies (e.g., lexical NPs). This is because, even though these policies tend to require greater constructive effort, they may be especially useful for increasing one's certainty that action will elicit cues to CG. Altogether, for a given motivational state, children are predicted to be less informative when certain and more informative when uncertain about CG.

The second relationship between deontic and pragmatic value speaks to the third set of findings (i.e., that relate repaired informativeness to feedback). Prior to feedback, children might tend to favor deontically valuable policies (Sacks & Schegloff, 1979). However, if children observe unexpected outcomes after initial reference (e.g., evidence of misunderstanding), their repaired reference might recruit pragmatically richer policies that require greater constructive effort. This sort of feedback-instigated (Pezzulo & Cisek, 2016), on-the-fly shift from deontic to pragmatic imperatives arguably accounts for the added informativeness of repaired compared to

initial reference (e.g., Nilsen & Mangal, 2012). Indeed, this shift should cut across speech act types. For example, imagine a child imperatively referring to an intended referent expecting their caregiver to retrieve it. Under the present model, children's ability to use referential control to elicit cues to CG should be key, and they should switch to costlier (repaired) referential control policies when such cues are not forthcoming. In line with this, Grosse et al. (2010) found that, compared to several control conditions, children less often repaired their imperative reference when caregivers correctly identified intended referents but did not retrieve them. Moreover, children often repaired their reference when caregivers "accidentally" retrieved intended referents (e.g., by handing over children's intended ball while saying *Oh, you want the paper?*). These results suggest that children's ability to elicit cues to CG is key, and that they will recruit costlier policies to guide repaired referential control when those expected sensations are not forthcoming.

What is the role of executive functioning in referential repair? Under the present model, when cues to CG (that indicate the source of referential failure) are absent from feedback, this adds demands on executive functioning that are not added when the source of failure is indicated. For example, when cues to CG are absent from feedback, children must infer the source of referential failure, maintain this inference in working memory, and use the contents of working memory as top-down context (i.e., contextual priors; Parr & Friston, 2017) for constructing and executing referential control policies. This example suggests that executive operations like working memory, task switching, and inhibitory control are implicated to a greater extent in repaired than initial reference; and in repaired reference following feedback with fewer than more cues to CG. One prediction that stems from this is that children's informativeness should suffer when feedback does not include cues to CG. Indeed, young children are more likely to simply repeat their initial reference, rather than construct more informative repairs, following feedback with fewer compared

to more cues to CG (Anselmi et al., 1986; Nilsen & Mangal, 2012). Nonetheless, evidence for the relationship of executive functioning to repaired informativeness from studies that manipulated feedback cues to CG is scant and provides mixed support for influences of executive functioning on production (e.g., Bacso & Nilsen, 2017; Nilsen & Graham, 2009; Wardlow & Heyman, 2016).

Conclusion. The preceding discussion can be summarized by responding to four challenges issued by Tison and Poirier (2021) (T&P) to the model of Vasil et al. (2020). First, T&P suggested that this model cannot account for complementary joint actions. However, three sets of findings that pertain to referential production – one half of a quintessentially complementary joint action – were accounted for using this model. The main idea is that young children expect to observe cues to CG and so are characteristically motivated to produce referential actions that efficiently elicit such cues. Most importantly, this gives rise to the inverse relationship between cues to CG and referential informativeness reviewed throughout this article. Second, T&P suggested that the model of Vasil et al. cannot account for imperative reference. This is incorrect. As has been seen, expectations of cues to CG and their effects on referential control cut across speech act types. Third, T&P suggest that the model presented here cannot account for how speakers manage cost-benefit tradeoffs as a "function of the context of utterance," (T&P, p. 7). However, it should be clear that young children's beliefs about context are central to the present model. The belief that one is in a collaborative context is predicted to motivate effortful control (i.e., that searches for and elicits cues to CG) whereas the belief that one is not in a collaborative context suppresses effortful control. Moreover, beliefs about the CG context bias competition among referential control policies, probabilistically favoring policies with epistemic, pragmatic, or deontic value[1].

Discussion

This section summarizes the review, discusses limitations, and proposes future directions.

Summary. This review argued for a positive view and against a deficits view (Glucksberg et al., 1975) of young children's referential informativeness. In support of this, experimental research on the relationship of cues to CG and children's informativeness was reviewed prior to elaborating the model of human communication that motivates the positive view (Vasil et al., 2020). Five points summarize the review. First, children's informativeness depends on cues to CG. When it comes to initial reference, informativeness possesses an inverse relationship with cues to CG whereas, for repaired reference, the relationship is a positive one. Second, young children's informativeness depends on participation in collaborative activities. Collaboration pulls for greater informativeness. Third, informativeness depends on linguistic prior to nonlinguistic cues to CG. Linguistic cues to CG influence informativeness by 2;0. In contrast, it is not until children are 2;6 that informativeness depends on nonlinguistic cues (see Figure 1). Though others have noted similar patterns (Moll & Kadipasaoglu, 2013), this is nonetheless surprising as, e.g., gestural reference depends on nonlinguistic cues to CG by around 12 months of age (Liszkowski, 2018) and children privilege nonlinguistic over linguistic reference in referential interpretation (Grassmann & Tomasello, 2010). Fourth, though evidence is shaky, cues to CG may interact with other cues to influence children's informativeness. Fifth, findings suggest a role for executive function in referential informativeness. Young children may have difficulty leveraging cues to CG when those cues invite conflicting perspectives on intended referents; and feedback that contains more cues to CG may positively influence repaired informativeness by easing cognitive demands.

Limitations. Two major limitations of the review are, first, that it focused on a very narrow slice of the experimental pie. This was due to the aims of the review, namely, to summarize literature on the causal relationship between cues to CG and young children's linguistic referential

informativeness. Though this focused perspective yielded conclusions summarized above, it neglected other research streams. For example, research on the effects of modeling suggests means by which children acquire strategies for reference (e.g., Carmiol et al., 2018; Matthews et al., 2007; Sarilar et al., 2015) and connects with their executive skills (e.g., Uzundag & Küntay, 2018). However, modeling manipulations do not implicate cues to CG about referents and so were not discussed. Additionally, research on cognitive processes that influence informativeness was not considered (e.g., attribute comparison; Davies & Kreysa, 2018; Rabagliati & Robertson, 2017).

Second, we did not discuss literature on multimodal reference. Indeed, this is an especially important oversight because linguistic reference emerges from (and, in the earliest stages, is even guided by) nonlinguistic reference (Tomasello, 2008; e.g., Iverson & Goldin-Meadow, 2005). Using observational methods, some authors report that multimodal reference depends on cues to CG at earlier ages than suggested by experimental studies of linguistic reference. Guerriero et al. (2006) reported that English and, to a lesser extent, Japanese speakers aged 1;5 often used points or touch alongside non-lexical (e.g., pronominal) reference to introduce new referents. Those authors concluded that infants understand that, when introducing referents, pronouns require additional (nonlinguistic) information for unambiguous reference whereas lexical reference does not. Ateş and Küntay (2018) and So et al. (2010) found related patterns among Turkish-speaking and English- and Chinese-speaking infants and toddlers. Generative models of the sort discussed in this article may be extended to any modality and may be multimodal (Parr et al., 2022).

Future directions. This review had four further limitations that, simultaneously, serve as suggested directions for research. First, this review collapsed across potentially important distinctions within types of cues to CG. For instance, variegated constructions were classified as examples of either more CG or less CG questions. Though this classification was useful, it is

conceivable that children's informativeness depends on finer distinctions (e.g., informativeness might depend more strongly on one type of "more CG" question than on others). Second, many of the studies reviewed above sampled only English-speaking children. This is an important caveat, as patterns of nominal realization vary along several dimensions (e.g., ergativity; Du Bois, 1987; argument drop; Valian, 1991). That is, while the major pattern observed here – namely, that informativeness tracks inversely with cues to CG – is applicable crosslinguistically as a kind of "universal" (e.g., Japanese; Guerriero et al., 2006; Turkish; Ateş & Küntay, 2018; Chinese; So et al., 2010; German; Salomo et al., 2011), each language offers children unique means to effect this inverse relationship (e.g., preferential use of null or pronominal forms; Guerriero et al., 2006; tripartite demonstrative systems; Küntay & Ozyürek, 2006). Third, we were limited in our ability to discern interactions between cues to CG with respect to their influence on informativeness. While there was some evidence for dependence across comparable studies, the lack of high-quality studies that crossed manipulations to nonlinguistic and linguistic cues to CG precludes certainty in these inferences. Fourth, only briefly discussed were effects of cues to cultural CG on children's informativeness; this was due to limited (nonlinguistic) or no (linguistic) research on the topic.

Conclusion. This article advanced a positive view on young children's linguistic referential informativeness. In support of the positive view, experimental research was reviewed that suggests that children's informativeness is inversely related to the presence of cues to CG, with dependence on linguistic cues to CG preceding dependence on nonlinguistic cues to CG. An active inference model of the internal processes that relate cues to CG to informativeness was elaborated.

Endnote

[1] Tison and Poirier's (2021) (T&P) issue a forth challenge. T&P suggest that "selective and targeted alignment of mental states relevant to… local goals" (*ibid*., p. 6) is important for joint action and that the model of Vasil et al. (2020) cannot account for this because it supposedly casts alignment as an "automatic process" (T&P, p. 6). The reasoning behind the latter claim is unclear. T&P seem to confuse Vasil et al.'s use of the word "alignment" with an unintended use of that word by Garrod and Pickering (e.g., 2009). However, Vasil et al. never cited that literature but, instead, cited literature on shared intentionality theory (e.g., Tomasello, 2008, 2014, 2019). In the latter, alignment relies on intentional (non-"automatic") communicative processes. Indeed, the adaptive prior for alignment is said to "foster intentional, patterned action…," (Vasil et al., 2020, p. 1).

Discussion of another challenge issued by T&P sheds further light on the adaptive prior for alignment. T&P seem to conflate synchronic and diachronic levels when they argue that Vasil et al. (2020) suggest "discontinuity" (T&P, p. 8) between human and nonhuman primate communication before labeling it a "saltationist view" (T&P, p. 8) of human evolution. Vasil et al. do argue for synchronic discontinuity, as this is an empirical fact (e.g., Bullinger et al., 2011; Duguid et al., 2014). However, diachronically, Vasil et al. suggest that human communicative motivations (captured by the adaptive prior for alignment) were "gradually cooperativized" (Vasil et al., 2020, p. 1) across human evolution. This is a non-saltationist claim. Indeed, the evolution of the adaptive prior for alignment may be modeled as gradual change in continuously valued quantities (see Figure 3). Others have suggested similar things, e.g., Bates et al. (1991) argued that "Variation along a continuous dimension can place the organism in a new "problem space," requiring qualitatively different solutions to ensure survival and reproductive success," (p. 34).

Acknowledgements

This article was based on a master's thesis written by the author. The author thanks Karl Friston, Rick Hoyle, Michael Tomasello, and Makeba Wilbourn, who were members on the author's master's thesis committee, for feedback on early versions of the manuscript. The author thanks Karl Friston, Anna Smith, Michael Tomasello, and anonymous reviewers for feedback on later versions of the manuscript.

References

Abbot-Smith, K., Nurmsoo, E., Croll, R., Ferguson, H., & Forrester, M. (2016). How children aged 2;6 tailor verbal expressions to interlocutor informational needs*. *Journal of Child Language*, *43*(6), 1277–1291. https://doi.org/10.1017/S0305000915000616

Anselmi, D., Tomasello, M., & Acunzo, M. (1986). Young children's responses to neutral and specific contingent queries. *Journal of Child Language*, *13*(1), 135–144. https://doi.org/10.1017/S0305000900000349

Ariel, M. (1988). Referring and Accessibility. *Journal of Linguistics*, *24*(1), 65–87.

Ateş, B., & Küntay, A. C. (2018). Referential interactions of Turkish-learning children with their caregivers about non-absent objects: Integration of non-verbal devices and prior discourse. *Journal of Child Language*, *45*(1), 148–173. https://doi.org/10.1017/S0305000917000150

Ateş-Sen, B., & Küntay, A. (2015). Children's sensitivity to caregiver cues and the role of adult feedback in the development of referential communication. In L. Serratrice & S. E. M. Allen (Eds.), *The Acquisition of Reference* (pp. 241–262). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.15.10ate

Bacso, S. A., & Nilsen, E. (2017). What's That You're Saying? Children With Better Executive Functioning Produce and Repair Communication More Effectively. *Journal of Cognition and Development*, *18*(4), 441–464. https://doi.org/10.1080/15248372.2017.1336438

Bacso, S. A., Nilsen, E., & Silva, J. (2021). How to turn that frown upside down: Children make use of a listener's facial cues to detect and (attempt to) repair miscommunication. *Journal of Experimental Child Psychology*, *207*, 105097. https://doi.org/10.1016/j.jecp.2021.105097

Badcock, P. B., Friston, K. J., Ramstead, M. J. D., Ploeger, A., & Hohwy, J. (2019). The hierarchically mechanistic mind: An evolutionary systems theory of the human brain, cognition, and behavior. *Cognitive, Affective & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-019-00721-3

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), 954–966.

Bates, E., Thai, D., & Marchman, V. (1991). Symbols and Syntax: A Darwinian Approach to Language Development. In N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch, & M. Studdert-Kennedy (Eds.), *Biological and Behavioral Determinants of Language Development* (pp. 29–65). Lawrence Erlbaum Associates. https://doi.org/10.4324/9781315807553-8

Berridge, K. C. (1996). Food reward: Brain substrates of wanting and liking. *Neuroscience and Biobehavioral Reviews*, *20*(1), 1–25. https://doi.org/10.1016/0149-7634(95)00033-b

Bjorklund, D. F., & Ellis, B. J. (2014). Children, childhood, and development in evolutionary perspective. *Developmental Review*, *34*(3), 225–264. https://doi.org/10.1016/j.dr.2014.05.005

Bohn, M., & Frank, M. C. (2019). The Pervasive Role of Pragmatics in Early Language. *Annual Review of Developmental Psychology*, *1*(1), 223–249. https://doi.org/10.1146/annurev-devpsych-121318-085037

Bohn, M., & Köymen, B. (2018). Common Ground and Development. *Child Development Perspectives*, *12*(2), 104–108.

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, *16*(10), 485–488. https://doi.org/10.1016/j.tics.2012.08.006

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482–1493. https://doi.org/10.1037/0278-7393.22.6.1482

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. *Synthese*, *195*(6), 2417–2444.

Bruner, J. S. (1983). *Child's Talk: Learning to Use Language*. Norton.

Bullinger, A. F., Zimmermann, F., Kaminski, J., & Tomasello, M. (2011). Different social motives in the gestural communication of chimpanzees and human children. *Developmental Science*, *14*(1), 58–68.

Bybee, J. (2002). Sequentiality as the basis of constituent structure. In *The Evolution of Language out of Pre-language* (pp. 107–134). John Benjamins. https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195301571.001.0001/acprof-9780195301571-chapter-15

Campbell, A. L., Brooks, P., & Tomasello, M. (2000). Factors Affecting Young Children's Use of Pronouns as Referring Expressions. *Journal of Speech, Language, and Hearing Research*, *43*(6), 1337–1349. https://doi.org/10.1044/jslhr.4306.1337

Carmiol, A. M., Matthews, D., & Rodríguez-Villagra, O. A. (2018). How children learn to produce appropriate referring expressions in narratives: The role of clarification requests

and modeling. *Journal of Child Language*, *45*(3), 736–752. https://doi.org/10.1017/S0305000917000381

Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, 159–181.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1–143.

Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1485), 1585–1599.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H., & Marshall, C. R. (1981). Definite Reference and Mutual Knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of Discourse Understanding* (pp. 10–63). Cambridge University Press.

Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, *2*(1), 19–41. https://doi.org/10.1080/01690968708406350

Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning & Verbal Behavior*, *22*(2), 245–258. https://doi.org/10.1016/S0022-5371(83)90189-5

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

Cohen, E. (2012). The Evolution of Tag-Based Cooperation in Humans: The Case for Accent. *Current Anthropology*, *53*(5), 588–616.

Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, *1*(2), 89–97.

Constant, A., Clark, A., & Friston, K. J. (2021). Representation Wars: Enacting an Armistice Through Active Inference. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.598733

Constant, A., Ramstead, M. J. D., Veissière, S. P. L., & Friston, K. (2019). Regimes of Expectations: An Active Inference Model of Social Conformity and Human Decision Making. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00679

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., & 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

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, *7*(5), 889–904. https://doi.org/10.1162/neco.1995.7.5.889

de Cat, C. (2015). The cognitive underpinnings of referential abilities. In L. Serratrice & S. E. M. Allen (Eds.), *The acquisition of reference* (pp. 263–283). John Benjamins.

Diesendruck, G. (2005). The Principles of Conventionality and Contrast in Word Learning: An Empirical Examination. *Developmental Psychology*, *41*(3), 451–463. https://doi.org/10.1037/0012-1649.41.3.451

Du Bois, J. W. (1987). The Discourse Basis of Ergativity. *Language*, *63*(4), 805–855. https://doi.org/10.2307/415719

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings. Biological Sciences / The Royal Society*, *281*(1796), 20141973.

Ford, W., & Olson, D. (1975). The elaboration of the noun phrase in children's description of objects. *Journal of Experimental Child Psychology*, *19*(3), 371–382. https://doi.org/10.1016/0022-0965(75)90068-5

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and Discourse Contributions to the Determination of Reference in Cross-Situational Word Learning. *Language Learning and Development*, *9*(1), 1–24. https://doi.org/10.1080/15475441.2012.707101

Frank, S. L., Bod, R., & Christiansen, M. H. (2012). How hierarchical is language use? *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20121741.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138.

Friston, K. (2012). A Free Energy Principle for Biological Systems. *Entropy*, *14*, 2100–2121.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active Inference: A Process Theory. *Neural Computation*, *29*(1), 1–49.

Friston, K., & Frith, C. (2015). A Duet for one. *Consciousness and Cognition*, *36*, 390–405.

Friston, K., Parr, T., Yufik, Y., Sajid, N., Price, C. J., & Holmes, E. (2020). Generative models, linguistic communication and active inference. *Neuroscience & Biobehavioral Reviews*, *118*, 42–64. https://doi.org/10.1016/j.neubiorev.2020.07.005

Friston, K., Rosch, R., Parr, T., Price, C., & Bowman, H. (2017). Deep temporal models and active inference. *Neuroscience and Biobehavioral Reviews*, *77*, 388–402.

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2014). The anatomy of choice: Dopamine and decision-making. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1655).

Garrod, S., & Pickering, M. J. (2009). Joint Action, Interactive Alignment, and Dialog. *Topics in Cognitive Science*, *1*(2), 292–304. https://doi.org/10.1111/j.1756-8765.2009.01020.x

Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, *30*(1), 5–56. https://doi.org/10.1515/ling.1992.30.1.5

Glucksberg, S., Krauss, R. M., & Higgins, E. T. (1975). The development of referential communication skills. In F. D. Horowitz (Ed.), *Review of Child Development Research* (Vol. 4). University of Chicago Press.

Glucksberg, S., Krauss, R. M., & Weisberg, R. (1966). Referential communication in nursery school children: Method and some preliminary findings. *Journal of Experimental Child Psychology*, *3*(4), 333–342. https://doi.org/10.1016/0022-0965(66)90077-4

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224.

Graf, E., & Davies, C. (2014). The production and comprehension of referring expressions. In D. Matthews (Ed.), *Pragmatic Development in First Language Acquisition* (pp. 161–182). John Benjamins.

Graf, E., Theakston, A., Lieven, E., & Tomasello, M. (2015). Subject and object omission in children's early transitive constructions: A discourse-pragmatic approach. *Applied Psycholinguistics*, *36*(3), 701–727. https://doi.org/10.1017/S0142716413000477

Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, *13*(1), 252–263. https://doi.org/10.1111/j.1467-7687.2009.00871.x

Greenfield, P. M., & Smith, J. H. (1976). *Structure of Communication in Early Language Development*. Academic Press. https://agris.fao.org/agris-search/search.do?recordID=US201300390435

Grigoroglou, M., & Papafragou, A. (2019a). Children's (and Adults') Production Adjustments to Generic and Particular Listener Needs. *Cognitive Science*, *43*(10), e12790. https://doi.org/10.1111/cogs.12790

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

Grosse, G., Behne, T., Carpenter, M., & Tomasello, M. (2010). Infants communicate in order to be understood. *Developmental Psychology*, *46*(6), 1710–1722. https://doi.org/10.1037/a0020727

Guerriero, A. M. S., Oshima-Takane, Y., & Kuriyama, Y. (2006). The development of referential choice in English and Japanese: A discourse-pragmatic perspective. *Journal of Child Language*, *33*(4), 823–857. https://doi.org/10.1017/s030500090600763x

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, *69*(2), 274–307. JSTOR. https://doi.org/10.2307/416535

Heins, C., Millidge, B., Demekas, D., Klein, B., Friston, K., Couzin, I. D., & Tschantz, A. (2022). pymdp: A Python library for active inference in discrete state spaces. *Journal of Open Source Software*, *7*(73), 4098. https://doi.org/10.21105/joss.04098

Hohwy, J. (2016). The self-evidencing brain. *Noûs*, *50*(2), 259–285.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367–371. https://doi.org/10.1111/j.0956-7976.2005.01542.x

Karimi, H. (2022). Greater entropy leads to more explicit referential forms during language production. *Cognition*, *225*, 105093. https://doi.org/10.1016/j.cognition.2022.105093

Kay, P. (1977). Language Evolution and Speech Style. In B. Blount & M. Sanches (Eds.), *Sociocultural Dimensions of Language Change* (pp. 21–33). Academic Press.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719.

Köymen, B., Rosenbaum, L., & Tomasello, M. (2014). Reasoning during joint decision-making by preschool peers. *Cognitive Development*, *32*, 74–85. https://doi.org/10.1016/j.cogdev.2014.09.001

Köymen, B., Schmerse, D., Lieven, E., & Tomasello, M. (2014). Young children create partner-specific referential pacts with peers. *Developmental Psychology*, *50*(10), 2334–2342. https://doi.org/10.1037/a0037837

Köymen, B., & Tomasello, M. (2020). The Early Ontogeny of Reason Giving. *Child Development Perspectives*, *14*(4), 215–220. https://doi.org/10.1111/cdep.12384

Küntay, A. C., & Ozyürek, A. (2006). Learning to use demonstratives in conversation: What do language specific strategies in Turkish reveal? *Journal of Child Language*, *33*(2), 303–320. https://doi.org/10.1017/s0305000906007380

Lambrecht, K. (1994). *Information structure and sentence form: Topics, focus, and the mental representations of discourse referents* (pp. xvi, 388). Cambridge University Press. https://doi.org/10.1017/CBO9780511620607

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical prerequisites: Vol. I*. Stanford University Press.

Levinson, S. C. (2006). On the human "interaction engine." In N. Enfield (Ed.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Berg.

Liebal, K., Carpenter, M., & Tomasello, M. (2013). Young children's understanding of cultural common ground. *The British Journal of Developmental Psychology*, *31*(Pt 1), 88–96.

Liszkowski, U. (2018). Origins and complexities of infant communication and social cognition. In A. Newen, L. De Bruin, & S. Gallagher (Eds.), *Oxford Handbook of 4E Cognition*. Oxford University Press. 10.1093/oxfordhb/97801998735410.013.35

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*(3), 297–307.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Reference and attitude in infant pointing. *Journal of Child Language*, *34*(1), 1–20.

Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two- and Four-Year-Olds Learn to Adapt Referring Expressions to Context: Effects of Distracters and Feedback on

Referential Communication. *Topics in Cognitive Science*, *4*, 184–210. 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

Mirza, M. B., Adams, R. A., Friston, K., & Parr, T. (2019). Introducing a Bayesian model of selective attention based on active inference. *Scientific Reports*, *9*(1), 1–22. https://doi.org/10.1038/s41598-019-50138-8

Moll, H., Carpenter, M., & Tomasello, M. (2011). Social Engagement Leads 2-Year-Olds to Overestimate Others' Knowledge. *Infancy*, *16*(3), 248–265. https://doi.org/10.1111/j.1532-7078.2010.00044.x

Moll, H., & Kadipasaoglu, D. (2013). The primacy of social over visual perspective-taking. *Frontiers in Human Neuroscience*, *7*. 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

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., & 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., & Graham, S. (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., & 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

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review*, *77*(4), 257–273. https://doi.org/10.1037/h0029436

O'Neill, D. K. (1996). Two-Year-Old Children's Sensitivity to a Parent's Knowledge State When Making Requests. *Child Development*, *67*(2), 659–677. https://doi.org/10.1111/j.1467-8624.1996.tb01758.x

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

O'Neill, G., & Miller, P. H. (2013). A show of hands: Relations between young children's gesturing and executive function. *Developmental Psychology*, *49*(8), 1517–1528. https://doi.org/10.1037/a0030241

Ortega, P. A., & Braun, D. A. (2013). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2153), 20120683.

Parr, T., & Friston, K. (2017). Working memory, attention, and salience in active inference. *Scientific Reports*, *7*(1), 14678.

Parr, T., Pezzulo, G., & Friston, K. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press.

Pea, R. D. (1979). Can information theory explain early word choice? *Journal of Child Language*, *6*(3), 397–410.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufmann.

Pezzulo, G., & Cisek, P. (2016). Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. *Trends in Cognitive Sciences*, *20*(6), 414–424. https://doi.org/10.1016/j.tics.2016.03.013

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35.

Pezzulo, G., Rigoli, F., & Friston, K. J. (2018). Hierarchical Active Inference: A Theory of Motivated Control. *Trends in Cognitive Sciences*, *22*(4), 294–306.

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances*, *6*(39), eabb5053. https://doi.org/10.1126/sciadv.abb5053

Prince, E. (1981). Toward a Taxonomy of Given-New Information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223–255).

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

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875–881.

Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2018). Answering Schrödinger's question: A free-energy formulation. *Physics of Life Reviews*, *24*, 1–16.

Ramstead, M. J. D., Veissière, S. P. L., & Kirmayer, L. J. (2016). Cultural Affordances: Scaffolding Local Worlds Through Shared Intentionality and Regimes of Attention. *Frontiers in Psychology*, *7*, 1090.

Rekers, Y., Haun, D. B. M., & Tomasello, M. (2011). Children, but not chimpanzees, prefer to collaborate. *Current Biology: CB*, *21*(20), 1756–1758.

Sacks, H., & Schegloff, E. A. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday Language* (pp. 15–21). Irvington.

Salomo, D., Graf, E., Lieven, E., & Tomasello, M. (2011). The role of perceptual availability and discourse context in young children's question answering. *Journal of Child Language*, *38*(4), 918–931. https://doi.org/10.1017/S0305000910000395

Salomo, D., Lieven, E., & Tomasello, M. (2010). Young Children's Sensitivity to New and Given Information when Answering Predicate-Focus Questions. *Applied Psycholinguistics*, *31*(1), 101–115. https://doi.org/10.1017/S014271640999018X

Salomo, D., Lieven, E., & Tomasello, M. (2013). Children's ability to answer different types of questions. *Journal of Child Language*, *40*(2), 469–491. https://doi.org/10.1017/S0305000912000050

Sarilar, A., Matthews, D., & Küntay, A. C. (2015). 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

Schegloff, E. A. (1972). Notes on a conversational practice: Formulating place. In D. N. Sudnow (Ed.), *Studies in Social Interaction* (pp. 75–119). MacMillan.

Schegloff, E. A. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. *Roots of Human Sociality: Culture, Cognition and Interaction*, 70–96.

Schiffer, S. R. (1972). *Meaning*. Oxford University Press.

Schmerse, D., Lieven, E., & Tomasello, M. (2015). Young children use shared experience to interpret definite reference. *Journal of Child Language*, *42*(5), 1146–1157. https://doi.org/10.1017/S0305000914000555

Schmidt, M. F. H., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young Children See a Single Action and Infer a Social Norm: Promiscuous Normativity in 3-Year-Olds. *Psychological Science*, *27*(10), 1360–1370.

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science*, *14*(3), 530–539.

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333.

Serratrice, L. (2008). The Role of Discourse and Perceptual Cues in the Choice of Referential Expressions in English Preschoolers, School-Age Children, and Adults. *Language Learning and Development*, *4*(4), 309–332. https://doi.org/10.1080/15475440802333619

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, *189*, 260–274.

Smith, R., Friston, K. J., & Whyte, C. J. (2022). A step-by-step tutorial on active inference and its application to empirical data. *Journal of Mathematical Psychology*, *107*, 102632. https://doi.org/10.1016/j.jmp.2021.102632

So, W. C., Demir, Ö. E., & Goldin-Meadow, S. (2010). When speech is ambiguous gesture steps in: Sensitivity to discourse-pragmatic principles in early childhood. *Applied Psycholinguistics*, *31*(1), 209–224. https://doi.org/10.1017/S0142716409990221

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Wiley.

Tison, R., & Poirier, P. (2021). Active Inference and Cooperative Communication: An Ecological Alternative to the Alignment View. *Frontiers in Psychology*, *12*, 3412. https://doi.org/10.3389/fpsyg.2021.708780

Tomasello, M. (1998). Reference: Intending that Others Jointly Attend. *Pragmatics and Cognition*, *6*, 229–243.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tomasello, M. (2014). *A Natural History of Human Thinking*. Harvard University Press.

Tomasello, M. (2018). How children come to understand false beliefs: A shared intentionality account. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(34), 8491–8498.

Tomasello, M. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard University Press.

Tomasello, M., Anselmi, D., & Farrar, M. J. (1984). Young children's coordination of gestural and linguistic reference. *First Language*, *5*(15), 199–209. https://doi.org/10.1177/014272378400501503

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

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1), 21–81. https://doi.org/10.1016/0010-0277(91)90046-7

Vasil, J. (accepted). The Study of Rational Framing Effects Needs Developmental Psychology [Commentary on: Rational Framing Effects: A Multidisciplinary Case by J. Bermúdez]. *Behavioral and Brain Sciences*.

Vasil, J., Badcock, P. B., Constant, A., Friston, K., & Ramstead, M. J. D. (2020). A World Unto Itself: Human Communication as Active Inference. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00417

Vasil, J., & Tomasello, M. (2022). Effects of "we"-framing on young children's commitment, sharing, and helping. *Journal of Experimental Child Psychology*, *214*, 105278. https://doi.org/10.1016/j.jecp.2021.105278

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, *43*. https://doi.org/10.1017/S0140525X19001213

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

Winters, J., Kirby, S., & Smith, K. (2018). Contextual predictability shapes signal autonomy. *Cognition*, *176*, 15–30.

Wittek, A., & Tomasello, M. (2005). Young children's sensitivity to listener knowledge and perceptual context in choosing referring expressions. *Applied Psycholinguistics*, *26*(4), 541–558. https://doi.org/10.1017/S0142716405050290

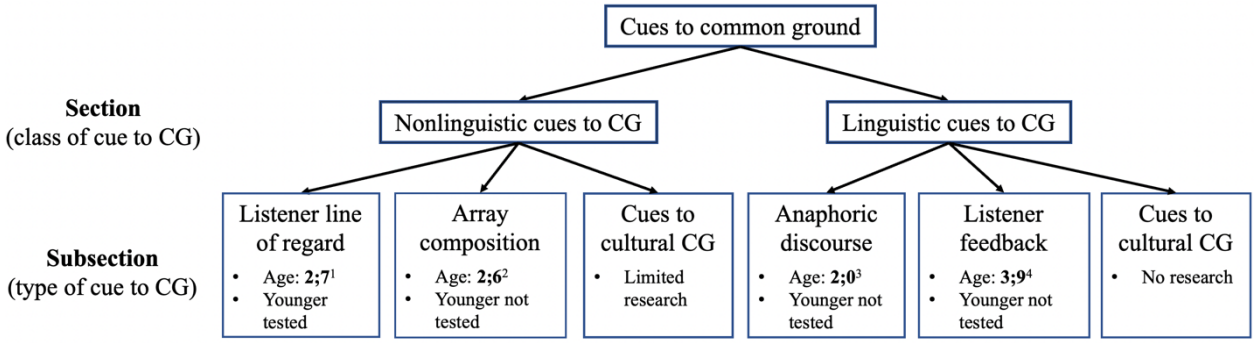

Figure 1. Roadmap for and summary of the literature reviewed in the second and third sections. The review follows from left to right. The first bullet point states the youngest reported age of group-level effects of the corresponding type of cue to CG on linguistic informativeness (mean sampled ages in cited studies; [1]O'Neill, 1996, Study 1; [2]Abbot-Smith et al., 2016; [3]Wittek & Tomasello, 2005, Study 2; [4]Nilsen & Mangal, 2012). The second bullet point states whether children younger than the age stated in the first bullet point have been tested.

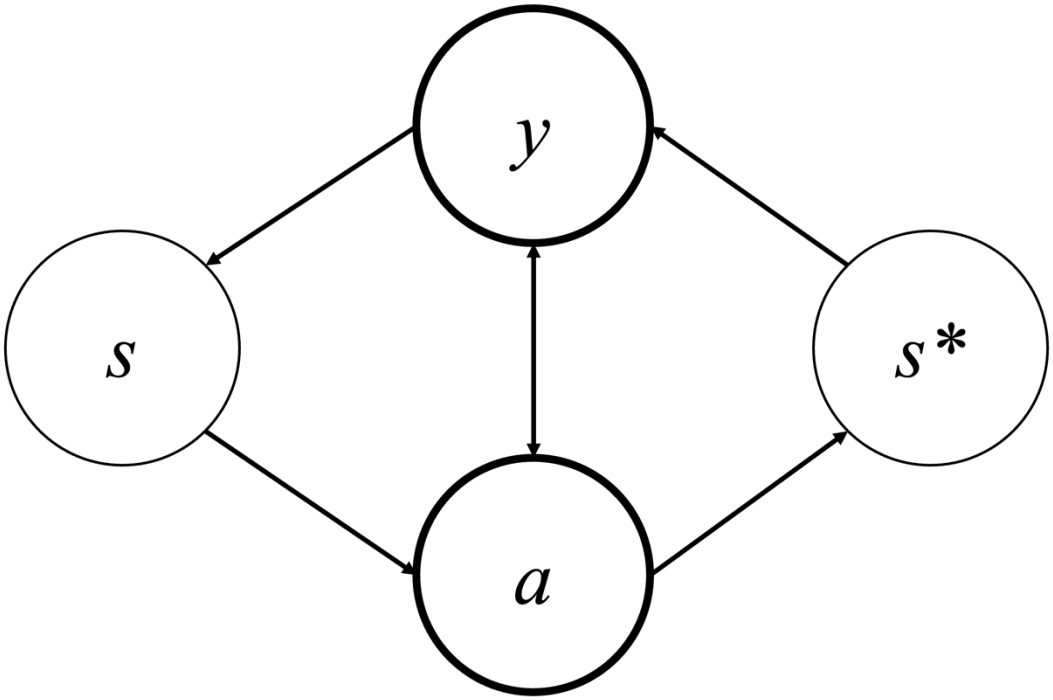

Figure 2. ***Markov blankets and circular causality.*** Markov blanket states in the bolded circles (i.e., sensory *y* and active *a* states) separate internal states *s* from external states *s**. Arrows indicate causal influences. Under this sparse causal coupling, internal states are conditionally independent of external states, and vice versa, given blanket states. This particular partition of states and the sparse coupling between these means that the dynamics of internal states parameterize conditional Bayesian beliefs (i.e., probability distributions) about the causes of sensory states (i.e., external and active states). The implicit neurocognitive message passing is specified by a generative model of how external and active states cause sensory states (see Figure 3).

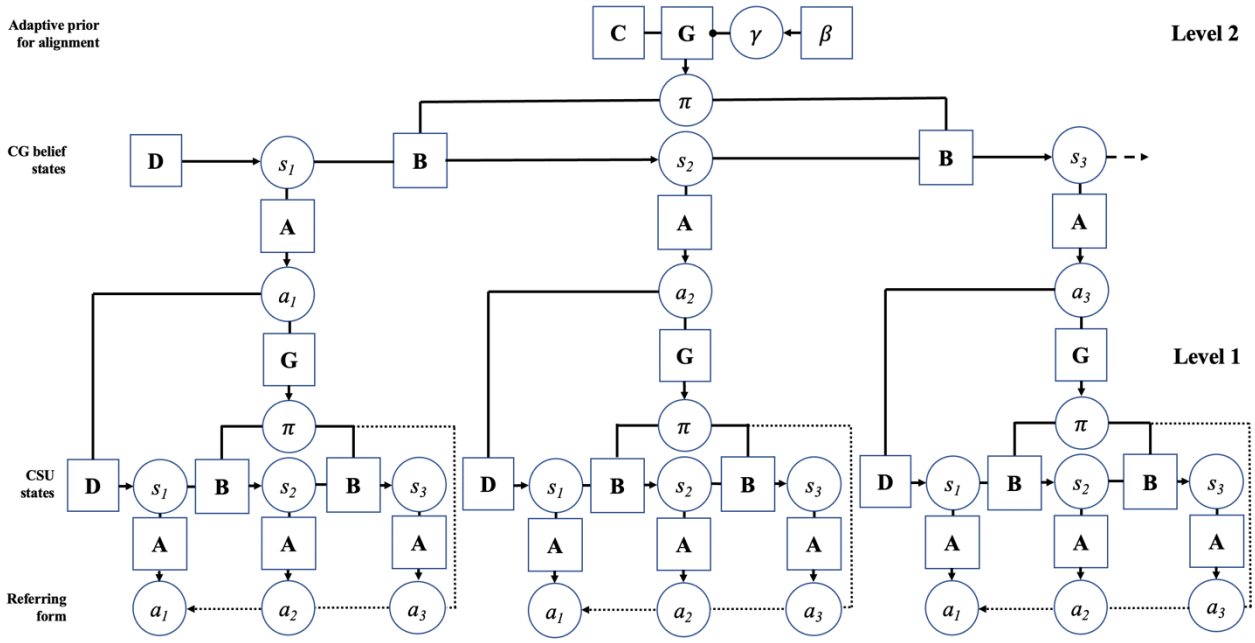

Figure 3. ***Hierarchical generative model for referential production.*** This figure casts the processes underpinning young children's referential informativeness as a partially observable Markov decision process (a common form of generative model for discrete states; Parr et al., 2022). Arrows indicate driving and nubs modulatory influences. In level 2, the adaptive prior for alignment specifies that young children have a strong prior belief (β) about the confidence or precision (γ) afforded the expected free energy (**G**) of policies (π) expected to elicit (*a priori* preferred) cues to CG ($\mathbf{C} = P(y|C)$) (see Friston et al., 2014). Policies guide transitions (**B**) between states *s* that encode beliefs about CG ($\mathbf{B} = P(s_\tau|s_{\tau-1}, \pi)$). Beliefs about CG provide top-down constraints on the expected free energy of competing referential control policies in level 1. Beliefs about CG favor either referential control policies that – starting from initial conditions $\mathbf{D} = P(s_1)$ – produce observable outcomes or actions *a* ($\mathbf{A} = P(a_\tau|s_\tau)$) that resolve uncertainty about CG (e.g., saccades that foveate to nonlinguistic cues to CG) or that elicit preferred outcomes (e.g., produce referring forms that elicit cues to CG indicative of mental state alignment). The latter includes policies with pragmatic (bold lines) or deontic affordance (dotted lines). Note that level

2 states convey beliefs to multiple level 1 states. This means that level 2 has slow dynamics that contextualize the fast dynamics of level 1, in the sense that 'what "we" are talking about' constrains 'what "I" or "you" are talking about.' There are numerous ways that this model may be expanded for different goals, e.g., placing priors over parameters (in squares and bolded font) to enable learning, adding precision to level 1, modifying hierarchical or temporal depth, etc. Candidate model structures may be used to simulate communication (Heins et al., 2022) or fitted to empirical responses to estimate the prior beliefs of individual subjects (Smith et al., 2022).