

How young children integrate information sources to infer the meaning of words

Manuel Bohn^{1,*}, Michael Henry Tessler^{2,*}, Megan Merrick³, & Michael C. Frank³

¹ Department of Comparative Cultural Psychology, Max Planck Institute for Evolutionary
Anthropology

² Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology

³ Department of Psychology, Stanford University

* These authors contributed equally to this work

Abstract

Before formal education begins, children typically acquire a vocabulary of thousands of words. This learning process requires the use of many different information sources in their social environment, including their current knowledge state as well as the context in which they hear words used. How is this information integrated? We specify a model according to which children consider all available information sources and integrate them via Bayesian inference. This model accurately predicted 2-to-5 year-old children's word learning across a range of experimental conditions. Model comparison suggests that the integration process is stable across development and that the main locus of development is an increased sensitivity to individual information sources. We present a quantitative developmental theory of information integration for language learning, embedded in a larger framework that grounds language use and learning in social cognition.

Keywords: language acquisition, social cognition, pragmatics, Bayesian modeling, common ground

Word count: X

How young children integrate information sources to infer the meaning of words

Humans rely on language for communication, making language acquisition a key learning objective of early childhood.¹ A central problem in language learning is referent identification: To learn the conventional symbolic relation between a word and an object, the child has to determine the intended referent of the word. Referents can only be identified in an inferential manner by reasoning about the speaker's intentions.²⁻⁵ That is, the child has to infer what the speaker is communicating about based on information sources in the social context of the utterance.

From early on in development, children use a number of different mechanisms to harness such social-contextual information sources.^{2,4,6} For example, children expect speakers to use novel words for unknown objects,⁷⁻¹⁰ to talk about objects that are relevant,^{11,12} new in context,^{13,14} or related to the ongoing conversation.¹⁵⁻¹⁷ These different mechanisms, however, have largely been described and theorized about in isolation. The picture of the learning process that emerges is that of a "bag of tricks": mechanisms that operate (and develop) independently from one another.⁶ As such, this view of the learning process does not capture the complexity of natural social interaction during which many sources of information are present.^{18,19} These information sources need to be integrated in order to accurately infer a speaker's intention.

When researchers do look at information integration in children,²⁰⁻²⁶ they mostly focus on how children interpret an ambiguous referring expression in light of social-contextual information. In one of these studies,²⁷ children faced a 2 x 2 display with a ball, a pen and two glasses in it. The speaker, sitting on the opposite side of the display, saw only 3 of the compartments: the ball, the pen, and one of the glasses. When the speaker asked for "the glass", children had to integrate the semantics of the utterance with the perspective of the speaker to correctly infer which of the glasses - from their perspective - the speaker was referring to. This study, along with others, found that preschoolers use

both information sources. What these studies do not do, however, is specify the process by which children integrate the different information sources. When discussing their results, authors refer to social-pragmatic theories of language use and learning,^{4,5,28–30} all of which assume that information is integrated as part of a social inference process, but none of which clearly defines the process. As a consequence, we have no explicit and quantitative theory of how different information sources (and word learning mechanisms) are integrated.

We present a theory of this integration process. Following social-pragmatic theories of language learning,^{4,5} our theory is based on the following premises: information sources serve different functional roles, all part of an integrated social inference process.^{28–31} Children use all available information to make inferences about the intentions behind a speaker’s utterance, which then leads them to correctly identify referents in the world and learn conventional word–object mappings. We formalize the computational steps that underlie this inference process in a cognitive model.^{32–34} This model makes quantitative predictions about how children should integrate information sources - predictions we can test against new data. It also allows us to test alternative theories about *whether* and *how* information is integrated. Finally, we can compare alternative views on how the integration process itself *develops*.

To test this model, we focus on three information sources that play an integral part in theorizing about language use and learning and operate on different timescales: (1) expectations that speakers communicate in a cooperative and informative manner,^{7,11} (2) shared common ground about what is being talked about in conversation,^{30,35} and (3) semantic knowledge about previously learned word–object mappings.^{6,36} While (1) is a momentary expectation about a particular utterance, (2) grows over the course of a conversation, and (3) is learned across development. This interplay of timescales has been hypothesized to be an important component of word meaning inference.^{37,38}

To test the predictive and explanatory power of our model, we developed an

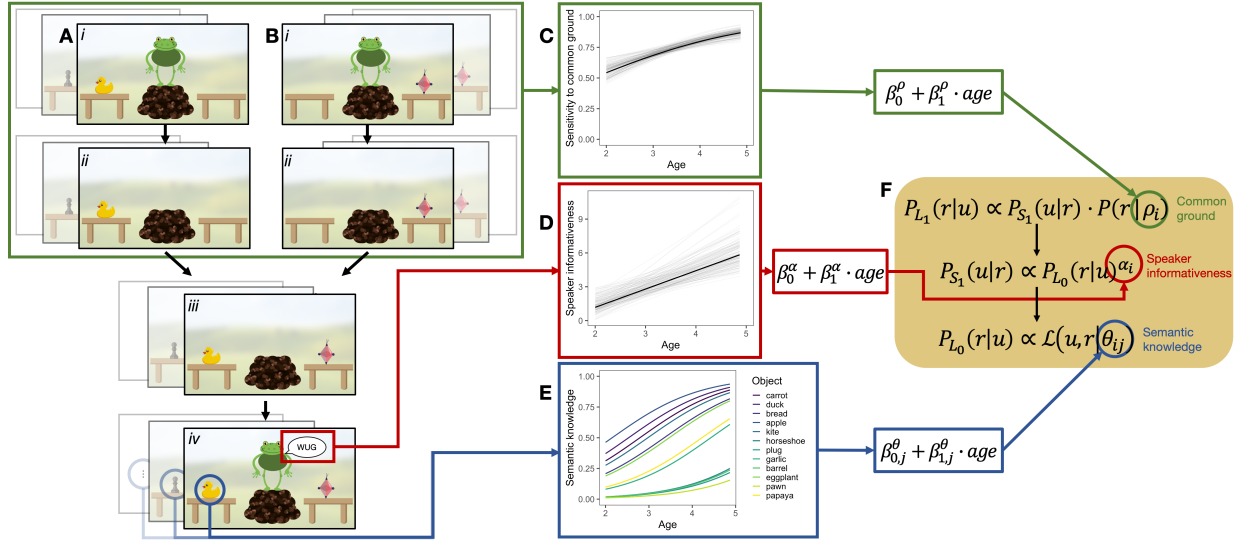


Figure 1. Experimental task and model. (A and B) Screenshots from the experimental task. (i) The speaker encounters one object and then leaves the scene (ii). While the speaker is away, a second object appears (iii), when returning, the speaker uses a novel word to request an object (iv). Sections (i) to (iii) establish common ground between the speaker and the listener, in that one object is new in context (green). The request in (iv) licenses an inference based on expectations about how informative speakers are (red). Listeners’ semantic knowledge enters the task because the identity of the known object on one of the tables is varied from well-known objects like a duck to relatively unfamiliar objects like a chess pawn (total of 12 objects - blue). (A) shows the condition in which common ground information is congruent (i.e. point to the same object) with speaker informativeness and (B) shows the incongruent condition. The congruent and incongruent conditions are each paired with the 12 known objects, resulting in 24 unique conditions. (C to E) show developmental trajectories for each information source, estimated from separate experiments (see supplement). (F) gives the model equation for the rational integration model and links information sources to model parameters.

world-learning paradigm in which we jointly manipulated these three information sources (Fig. 1). Children interacted on a tablet computer with a series of storybook speakers.³⁹ The speakers used novel words (e.g., “wug”) in the context of two potential referents. This situation is depicted in Fig. 1A-iv, in which a speaker (here, a frog) appears with a known object (a duck, left) and an unfamiliar object (the diamond-shaped object, right). To identify the referent, the child could rely upon the three information sources described above. First, they may infer that a cooperative speaker would have used the word duck to refer to the known object (the duck); the fact that the speaker did not say duck then

suggests that the speaker is most likely referring to a different object (the unfamiliar object). Thus, the listener can infer the referent of the novel word by assuming that the speaker is a rational agent who chooses the most informative words to refer to objects.¹¹ Second, the child may rely upon what has already been established in the common ground with the speaker. Listeners expect speakers to communicate about things that are new to the common ground.^{13,14} Thus, the inference about the novel word referring to the unfamiliar object also depends on which object is new in context (Fig. 1A and B i-iii). Finally, the child may rely upon their previously acquired semantic knowledge, that is, how sure they are that the known object is called “duck”. If the known object is something less familiar, such as a chess piece (e.g., a pawn), a 3-year-old child may draw a weaker inference, if they draw any inference at all.⁴⁰ Taken together, the child has the opportunity to integrate their assumptions about (1) informative communication, (2) their understanding of the common ground, and (3) their existing semantic knowledge. In some versions of the experiment, some of these information sources were aligned (Fig. 1A) but in some they were in conflict (Fig. 1B).

Our *rational integration model* arbitrates between information sources via Bayesian inference (see Fig. 1F for model formula). A listener (L_1) reasons about the referent of a speaker’s (S_1) utterance. This reasoning is contextualized by the prior probability of each referent ρ which we take to be a function of the common ground shared between the listener and the speaker. To decide between referents, the listener (L_1) reasons about what a rational speaker (S_1) with informativeness α would say given an intended referent. This speaker is assumed to compute the informativity for each available utterance and then choose the most informative one. The informativity of each utterance is given by imagining which referent a listener who interprets words according to their literal semantics (what we call a literal listener, L_0), would infer upon hearing the utterance. Naturally, this reasoning depends on what kind of semantic knowledge (for object j) θ_j the speaker ascribes to the (literal) listener. Taken together, this model provides a quantitative theory of information

integration during language learning. Furthermore, the model presents an explicit and substantive theory of development. It assumes that, while children’s sensitivity to the individual information sources increases with age, the way integration proceeds remains constant.^{2,41} In the model, this is accomplished by making the parameters capturing sensitivity to speaker informativeness (α_i , Fig. 1D), the common ground (ρ_i , Fig. 1C), and object specific semantic knowledge ($\theta_{i,j}$, Fig. 1E) a function of the child’s age.

Results

Predicting information integration accross development

We tested the model in its ability to predict children’s actual behavior - at an age (between 2 and 5 years) at which they learn language at an exponential rate.⁶ We estimated children’s (N=148) developing sensitivity to individual information sources in separate experiments (see Fig. 1C-E). We then generated parameter-free, quantitative, *a priori* model predictions (developmental trajectories) representing the model’s expectations about how children should behave in a new situation in which all three information sources had to be integrated. We generated model predictions for 24 experimental conditions: 12 objects of different familiarities (requiring different levels of semantic knowledge), with novelty either conflicting or coinciding; Fig. 1. We compared these predictions to newly collected data from N = 220 children from the same age range. The results showed a very close alignment between model predictions and the data across the entire age range. That is, the average developmental trajectories predicted by the model resembled the trajectories found in the data (Fig. S6). For a more quantitative analysis, we binned predictions and data by year and correlated them. We found a high correlation showing that the model explained 79% of the variance in the data (Fig. 2). These results seem to support the assumption in the model that children consider all available information source. However, it is still possible that simpler models might make equally good or better predictions.

Thus, we formalized the alternative view that children selectively ignore information sources in the form of three lesioned models (Fig. 2). These models assume that children follow the heuristic “ignore X” (with X being one of the information sources) when multiple information sources are presented together.

The marginally less-complex *no word knowledge model* uses the same model architecture as the rational integration model. It takes in expectations about speaker informativeness and common ground, but omits semantic knowledge that is specific to the familiar objects (i.e., uses only general semantic knowledge). The object specific parameters for semantic knowledge are fitted via a hierarchical regression (mixed effects) model. In this model, there is an overall developmental trajectory for semantic knowledge (main effect) and then there is object-specific variation around this trajectory (random effects). The *no word knowledge model* takes in the overall trajectory for semantic knowledge but ignores object-specific variation. That is, the model assumes a listener whose inference does not vary depending on the particular familiar object but only depends on the age specific average semantic knowledge. The *no common ground model* takes in object specific semantic knowledge and speaker informativeness but ignores common ground information. Instead of assuming that one object has a higher prior probability to be the referent because it is new in context, the speaker thinks that both objects are equally likely to be the referent. As a consequence, the listener does not differentiate between situations in which common ground is aligned or in conflict with the other information sources. Finally, according to the *no speaker informativeness model*, the listener does not assume that the speaker is communicating in an informative way and hence ignores the information provided by the utterance. As a consequence, the inference is solely based on common ground expectations.

We found little support that children make inferences following one of the heuristic models. When using Bayesian model comparison via marginal likelihood of the data,⁴² we find that the data was several orders of magnitude more likely under the rational integration model compared to any of the lesioned models (Fig. 2). Taken together, we

162 may conclude that children considered all available information sources.

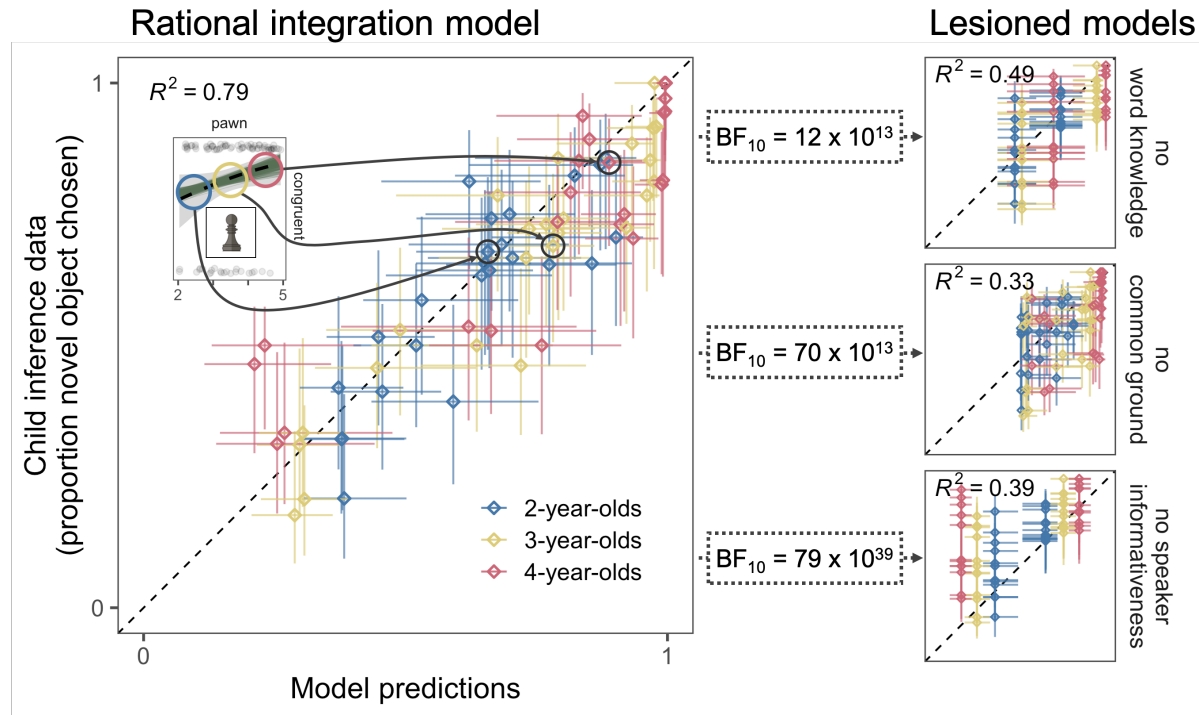


Figure 2. Predicting information integration. Correlation between model predictions and child inference data for all 24 conditions and for each age group (binned by year) for the rational integration model and the three lesioned models. Horizontal and vertical error bars show 95% HDI. Inset shows an example of model predictions as developmental trajectories (see Fig. 3). The no word knowledge model includes developmental change for overall vocabulary development but ignores the identity of the known object. The no common ground model assumes that children ignore information from common ground and the no speaker informativeness model assumes that children ignore inferences that are licensed by expectations about informative communication. BF_{10} gives the Bayes Factor in favor of the integration model based on the marginal likelihood of the data under each model.

163 Explaining the process of information integration

164 The fact that children integrated all the available information sources may be
 165 surprising, but does not rule out that the integration process itself might be different from
 166 the one described in the rational integration model. Remember that this model assumes
 167 that all information sources enter into an integrated inference process. However, a more
 168 modular view seems equally plausible: children might compute separate inferences based on

a subset of the available information and then integrate them by weighting them according to some ratio. This view would allow for the possibility that some information sources are considered to be more important than others. In other words, children might be biased towards some information sources (e.g. a 60:40 ratio). We formalized this alternative view as a *biased integration model*. This model assumes that semantic knowledge and expectations about speaker informativeness enter into one inference (mutual exclusivity inference^{7,8,40}) while common ground information enters into a second one. The outcome of both processes are then weighted according to the parameter ϕ . Like the rational integration model, this model takes in all available integration sources and assumes that they are integrated, the integration process itself, however, is construed in a different way.

The parameter ϕ in the biased integration model is unknown ahead of time and has to be estimated based on the experimental data. That is, in contrast to the rational integration model, the biased integration model does not allow to make *a priori* predictions about new data in the way we describe above. For the model comparison, we therefore constrained the parameters in the rational integration model by the experimental data as well. As a consequence, both models use all the information available to constrain their parameters (see Fig. S4).

The biased integration model makes reasonable predictions that explain a large proportion of the variance in the data (Fig. 3B). The parameter ϕ - indicating the bias to one of the inferences - was estimated to favor the mutual exclusivity inference (mode = 0.65; 95% HDI: 0.60 - 0.71, see Fig. 3B). However, the rational integration model presented a much better fit to the data, both in terms of correlation and marginal likelihood of the data (Fig. 3B). Thus, a fully integrated - as opposed to a segmented, biased - integration process explains the data better.

The rational integration model assumes that the integration process itself does not change with age.² That is, while children's sensitivity to each information source develops,

the way they relate to one another remains the same. An alternative perspective would be to assume that the integration process itself changes with age. The biased integration model provides a way to think about developmental change in the integration process: Children are biased towards some information sources, but this bias itself changes with age. We formalize such an alternative view as a *developmental bias model* which is structurally identical to the biased integration model but in which the parameter ϕ changes with age. The model assumes that the importance of the different information sources changes with age.

The developmental bias model explains a substantial portion of the variance in the data (Fig. 3B). The estimated developmental trajectory for the bias parameter ϕ (Fig. 3B) suggests that younger children put a higher emphasis on common ground information, while older children rely more on the mutual exclusivity inference. Yet again, when we directly compare the competitor models, we find that the data is several orders of magnitude more likely under the rational integration model (Fig. 3B). Interestingly, when we compare the rational integration model with its parameters not constrained by the experimental data (i.e. the “pure” prediction model described above), it also outperformed both bias models (see Fig. S10 and Tab. S6). Taken together, these results suggest that the integration process is best described as a) fully integrated and b) stable over time.

Discussion

These findings show that young children can flexibly integrate multiple information sources during language learning, even from relatively early in development. To answer the question of how they do so, we presented a formal cognitive model that assumes that information sources are rationally integrated via Bayesian inference. This model accurately predicted children’s behavior between 2 and 5 years of age and provided a better explanation of the integration process and its development compared to a range of competitor models that either neglected or prioritized some information sources.

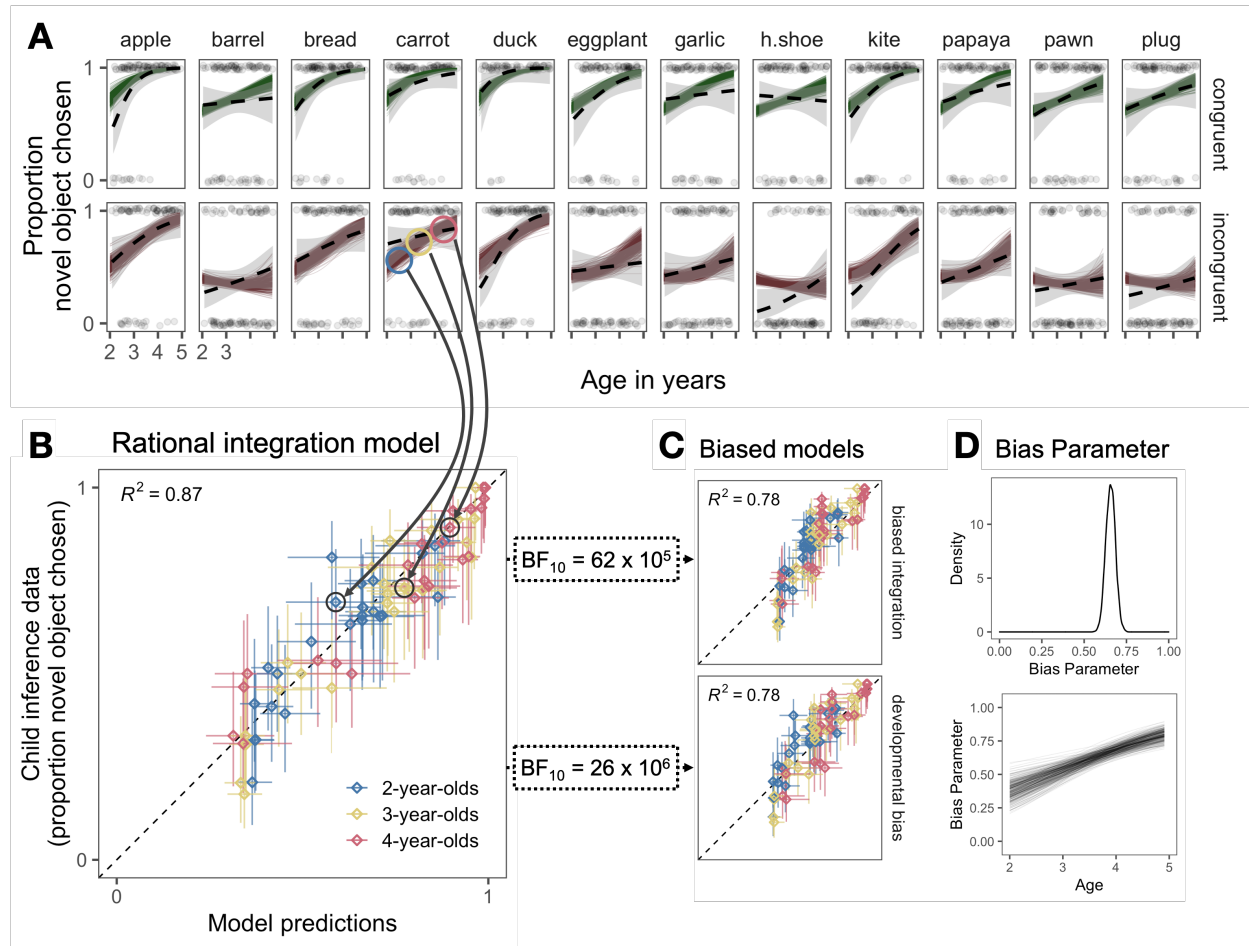


Figure 3. Explaining information integration across development. (A) Model predictions from the rational integration model (colored lines) next to the behavioral data (dotted black lines with 95% CI in gray) for all 24 experimental conditions. Top row (blue) shows conditions in which common ground information is congruent with speaker informativeness, bottom row (red) shows conditions in which they are incongruent. Known objects are ordered based on their rated age of acquisition (left to right). Light dots represent individual data points. (B) Correlations between model predictions binned by age and condition for the integration model and (C) the two mixture models. Vertical and horizontal error bars show 95% HDIs. BF_{10} gives the Bayes Factor in favor of the rational integration model based on the marginal likelihood of the data under each model. (D) Posterior distribution of the bias parameter in the biased integration model and 300 random draws from the posterior distribution of the intercept and slope for the bias parameter in the developmental bias model.

Replicate with other languages and other cultural settings.

more explicit model of common ground (see perspective taking work mcc)

integrate process of word learning itself via cross-situational learning^{37,38} dan

yurovski work

The model is derived from a more general framework for pragmatic inference, which has been used to explain a wide variety of phenomena in adults' language use and comprehension.^{43–46} Thus, it can be generalized in a natural way to capture word learning in contexts that offer more, fewer, or different types of information.⁴¹ The flexibility of the framework stems from its conceptualization of human communication as a form of rational social action.

Connections to naive utility calculus stuff?

As such, this work adds to theories that see the onto- and phylogenetic emergence of language as deeply rooted in social cognition.

Methods

Participants

Materials

Procedure

Data analysis

References

1. Tomasello, M. *A natural history of human thinking*. (Harvard University Press, 2018).
2. Bohn, M. & Frank, M. C. The pervasive role of pragmatics in early language. *Annual Review of Developmental Psychology* **1**, 223–249 (2019).
3. Bruner, J. *Child's talk: Learning to use language*. (New York: Norton, 1983).
4. Clark, E. V. *First language acquisition*. (Cambridge: Cambridge University Press, 2009).
5. Tomasello, M. *Constructing a language*. (Cambridge, MA: Harvard University Press, 2009).
6. Bloom, P. *How children learn the meanings of words*. (MIT press, 2002).
7. Clark, E. V. The principle of contrast: A constraint on language acquisition. *Mechanisms of language acquisition* (1987).
8. Markman, E. M. & Wachtel, G. F. Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology* **20**, 121–157 (1988).
9. Carey, S. & Bartlett, E. Acquiring a single new word. *Proceedings of the Stanford Child Language Conference* **15**, 17–29 (1978).
10. Halberda, J. The development of a word-learning strategy. *Cognition* **87**, B23–B34 (2003).
11. Frank, M. C. & Goodman, N. D. Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology* **75**, 80–96 (2014).
12. Schulze, C., Grassmann, S. & Tomasello, M. 3-year-old children make relevance inferences in indirect verbal communication. *Child Development* **84**, 2079–2093 (2013).
13. Akhtar, N., Carpenter, M. & Tomasello, M. The role of discourse novelty in early word learning. *Child Development* **67**, 635–645 (1996).

14. Diesendruck, G., Markson, L., Akhtar, N. & Reudor, A. Two-year-olds' sensitivity to speakers' intent: An alternative account of samuelson and smith. *Developmental Science* **7**, 33–41 (2004).
15. Bohn, M., Le, K., Peloquin, B., Koymen, B. & Frank, M. C. Children's interpretation of ambiguous pronouns based on prior discourse. *PsyArXiv* (2020). doi:10.31234/osf.io/gkhez
16. Sullivan, J. & Barner, D. Discourse bootstrapping: Preschoolers use linguistic discourse to learn new words. *Developmental Science* **19**, 63–75 (2016).
17. Horowitz, A. C. & Frank, M. C. Young children's developing sensitivity to discourse continuity as a cue for inferring reference. *Journal of experimental child psychology* **129**, 84–97 (2015).
18. Bergelson, E. & Aslin, R. N. Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences* **114**, 12916–12921 (2017).
19. Frank, M. C., Tenenbaum, J. B. & Fernald, A. Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development* **9**, 1–24 (2013).
20. Ganea, P. A. & Saylor, M. M. Infants' use of shared linguistic information to clarify ambiguous requests. *Child Development* **78**, 493–502 (2007).
21. Graham, S. A., San Juan, V. & Khu, M. Words are not enough: How preschoolers' integration of perspective and emotion informs their referential understanding. *Journal of Child Language* **44**, 500–526 (2017).
22. Grosse, G., Moll, H. & Tomasello, M. 21-month-olds understand the cooperative logic of requests. *Journal of Pragmatics* **42**, 3377–3383 (2010).
23. Jara-Ettinger, J., Floyd, S., Huey, H., Tenenbaum, J. B. & Schulz, L. E. Social pragmatics: Preschoolers rely on commonsense psychology to resolve referential

underspecification. *Child Development* **91**, 1135–1149 (2020).

24. Khu, M., Chambers, C. G. & Graham, S. A. Preschoolers flexibly shift between speakers' perspectives during real-time language comprehension. *Child development* **91**, e619–e634 (2020).

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

26. Nilsen, E. S., Graham, S. A. & Pettigrew, T. Preschoolers' word mapping: The interplay between labelling context and specificity of speaker information. *Journal of child language* **36**, 673–684 (2009).

27. Nadig, A. S. & Sedivy, J. C. Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science* **13**, 329–336 (2002).

28. Grice, H. P. *Studies in the way of words*. (Cambridge, MA: Harvard University Press, 1991).

29. Sperber, D. & Wilson, D. *Relevance: Communication and cognition*. (Cambridge, MA: Blackwell Publishers, 2001).

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

31. Shafto, P., Goodman, N. D. & Frank, M. C. Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science* **7**, 341–351 (2012).

32. Frank, M. C. & Goodman, N. D. Predicting pragmatic reasoning in language games. *Science* **336**, 998–998 (2012).

33. Goodman, N. D. & Frank, M. C. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences* **20**, 818–829 (2016).

34. Oberauer, K. & Lewandowsky, S. Addressing the theory crisis in psychology.
Psychonomic bulletin & review **26**, 1596–1618 (2019).
35. Bohn, M. & Koymen, B. Common ground and development. *Child Development Perspectives* **12**, 104–108 (2018).
36. Fenson, L. *et al.* Variability in early communicative development. *Monographs of the Society for Research in Child Development* i–185 (1994).
37. Frank, M. C., Goodman, N. D. & Tenenbaum, J. B. Using speakers’ referential intentions to model early cross-situational word learning. *Psychological Science* **20**, 578–585 (2009).
38. McMurray, B., Horst, J. S. & Samuelson, L. K. Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review* **119**, 831 (2012).
39. Frank, M. C., Sugarman, E., Horowitz, A. C., Lewis, M. L. & Yurovsky, D. Using tablets to collect data from young children. *Journal of Cognition and Development* **17**, 1–17 (2016).
40. Lewis, M. L., Cristiano, V., Lake, B. M., Kwan, T. & Frank, M. C. The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition* **198**, 104191 (2020).
41. Bohn, M., Tessler, M. H., Merrick, M. & Frank, M. C. Predicting pragmatic cue integration in adults’ and children’s inferences about novel word meanings. (2019).
doi:10.31234/osf.io/xma4f
42. Lee, M. D. & Wagenmakers, E.-J. *Bayesian cognitive modeling: A practical course*. (Cambridge University Press, 2014).
43. Monroe, W., Hawkins, R. X., Goodman, N. D. & Potts, C. Colors in context: A pragmatic neural model for grounded language understanding. *Transactions of the*

337 *Association for Computational Linguistics* **5**, 325–338 (2017).

338 44. Wang, S., Liang, P. & Manning, C. D. Learning language games through interaction.

339 *Proceedings of the 54th Annual Meeting of the Association for Computational*

340 *Linguistics* 2368–2378 (2016).

341 45. Tessler, M. H. & Goodman, N. D. The language of generalization. *Psychological Review*

342 **126**, 395–436 (2019).

343 46. Yoon, E. J., Frank, M. C., Tessler, M. H. & Goodman, N. D. Polite speech emerges

344 from competing social goals. *psyarxiv.com/67ne8* (2018).