Broke or caused to break: How children map causal verbs to different causes across development

David Rose[1], Siying Zhang[2], Shaun Nichols[3], Ellen M. Markman[1], and Tobias Gerstenberg[1]

[1]Department of Psychology. Stanford University

[2]Department of Psychology, University of Washington

[3]Department of Philosophy, Cornell University

Abstract

Although collision-like causes are fundamental in philosophical and psychological theories of causation, humans conceptualize many events as causes that lack direct contact. We argue that how people think and talk about different causes is deeply connected and investigate how children learn this mapping. If Andy hits Suzy with his bike, Suzy falls into a fence and it breaks, Andy *caused* the fence to break but Suzy *broke* it. If Suzy forgets sunscreen and gets sunburned, the absence of sunscreen *caused* Suzy's sunburn, but the sun *burned* her skin. We tested 691 children and 150 adults. Four-year-old children mapped "caused" to distal causes and "broke" to proximal causes (Experiment 1). Though four-year-old children didn't map "caused" to absences until later (Experiment 2), they already referred to absences when asked "why" an outcome occurred (Experiment 3). Our findings highlight the role of semantics and pragmatics in developing these mappings.

Keywords: causation; language understanding; semantics; pragmatics; cognitive development

---

Corresponding author: David Rose (davdrose@stanford.edu). All the data, study materials, pre-registrations, and analysis code are available here: `https://github.com/davdrose/cause_burn_development`

Introduction

Some say that causation is the cement of the universe [1]. Others say that it's the glue between observed events [2]. Whatever your views on adhesion, determining what causes what underlies many of the most profound cultural and scientific achievements, from the discovery of fire to the discovery of germs. The capacity for causal reasoning is a fundamental building block of cognition, with some aspects of causal cognition being unique to humans [e.g., 3], including that we use different causal expressions to refer to different types of causes [e.g., 4].

Linguists distinguish between lexical and periphrastic causatives. Lexical causatives encode causal relationships in a single verb – such as "break", "burn", and "crack". Periphrastic causatives are two-clause expressions, such as "caused to break" [e.g., 5]. Lexical causatives can refer only to direct causes while periphrastic causatives can also refer to indirect causes [6, 7, 8, 9, 10, 11, 12, 5]. Linguists usually characterize direct causes as those that contact their effects [e.g., 9, 13, 14, 15, 16, 8, see also 12, for discussion]. For instance, in determining what caused what when Andy hits Suzy with his bike, Suzy falls into a fence, and the fence breaks, it seems appropriate to say that Andy *caused* the fence to break and that Suzy *broke* it [12, 4].

The linguistic distinction between lexical and periphrastic causatives relates to the philosophical distinction between productive and dependence-based causes [17, 18]. Productive causes bring about effects via a spatiotemporally continuous process [19, 20, 12]. This can involve direct contact between a single cause and effect—such as Suzy kicking a fence and breaking it—but can also involve chains of events. In the fence-breaking scenario where Andy hits Suzy with his bike, she falls into the fence and it breaks, the distal cause, Andy, initiates the chain and the proximal cause, Suzy, makes contact with the effect. Dependence-based causes can be defined in terms of counterfactuals, such as "if the cause hadn't occurred, then the effect wouldn't have occurred" [21]. Accordingly, absences can also be causes. Not watering plants can cause them to die. Here, there is no production but there is dependence: if the plants had been watered, they wouldn't have died [22, 23, 24, 25, 26, 27].

Rose, Sievers and Nichols [4] show that these linguistic and philosophical distinctions are reflected in adults' language use. For example, if Suzy goes to the beach, forgets her sunscreen and gets a sunburn, adults judge that the sun *burned* her skin and that the absence of sunscreen *caused her skin to burn.* Rose et al [4] suggest that lexical causatives, such as "burn" tend to be used for productive causes, and that "caused" tends to be used for non-productive causes, such as absences.

Here, we investigate how children learn to map different causal expressions to different causes. To do so, they need to master two challenges: a semantic challenge of learning the meaning of different causal expressions, and a pragmatic challenge of understanding what expression is most appropriate to use in context. The semantic challenge includes realizing that lexical causatives have a narrow domain of application, whereas periphrastic causatives apply more broadly. For example, it is literally true that in the fence-breaking scenario, both Andy and Suzy *caused* the fence to break. But only Suzy *broke* it. Children need to learn that "caused" can refer not only to direct causes but also to indirect causes and absences. The semantic overlap between lexical and periphrastic causatives creates the pragmatic challenge. When someone asks "Who caused the fence to break?", it's more appropriate to say that it was Andy (rather than Suzy). Suzy not only caused the fence to break, she broke it. Because there is a more specific expression that one could have used to refer to Suzy, the less specific expression points to Andy. By examining how children map causal expressions to causes across development, we gain insight into how they address the semantic and pragmatic challenge, revealing the psychological mechanisms underlying the connection between the causal expressions that linguists recognize and the causal relations that philosophers characterize. Before turning to our experiments, we briefly discuss prior work on how children's causal cognition and causal language develops.

Development of causal cognition

Direct and indirect causes

Direct contact matters for how infants perceive causation. Infants perceive causation when one object makes contact with another, and stops moving at the same time as the other one starts to move. If there is no contact, or a temporal delay, but the second object moves anyhow, infants don't perceive causation [28, 29, 30].

Direct contact—though not the only thing that matters—also plays a role in how children reason about causation [31]. When two-year-old children are shown that a toy airplane lights up when a block comes into contact with its base, they will touch the base with the block when asked to make it go. In contrast, when the airplane is connected to the base by a long wire, they don't touch the base with the block when asked to make it go [32]. Children appear to understand direct causal relationships earlier than indirect ones.

Absent causes

Around the age of 3 to 4, children grasp that causation doesn't always require contact. For instance, they will hold an object over a detector (without touching it) after seeing that doing so makes the detector light up [32, 33, 34, 35, 36]. Around 5 years of age, children recognize some absences (e.g., not salting an icy sidewalk) as causes (e.g., someone falling) [37]. Because absences don't produce effects, treating absences as causes suggests an understanding of dependence-based causation.

Development of causal language

Lexical and periphrastic causatives

Lexical causatives, such as "break", "drop", "dry", and "open", are included in the vocabulary of most English-speaking children between the ages of 2 and 2½ [38]. Around this time, children use them in adult-like ways both in their transitive form (e.g., "Andy broke the fence.") and their intransitive form (e.g., "The fence broke.") [39, 40].

At this same age, children understand novel verbs used in transitive, but not intransitive sentences, as having causal meaning. Naigles [41] showed 2-year-old children scenes where two actions were performed at the same time. In the causal scene, a bunny pushed down on a duck's head, making the duck tilt its head. In the non-causal scene, the duck and bunny twirled their arms in synchrony. Children heard novel verbs in either transitive form ("The duck is krading the bunny.") or intransitive form ("The duck and the bunny

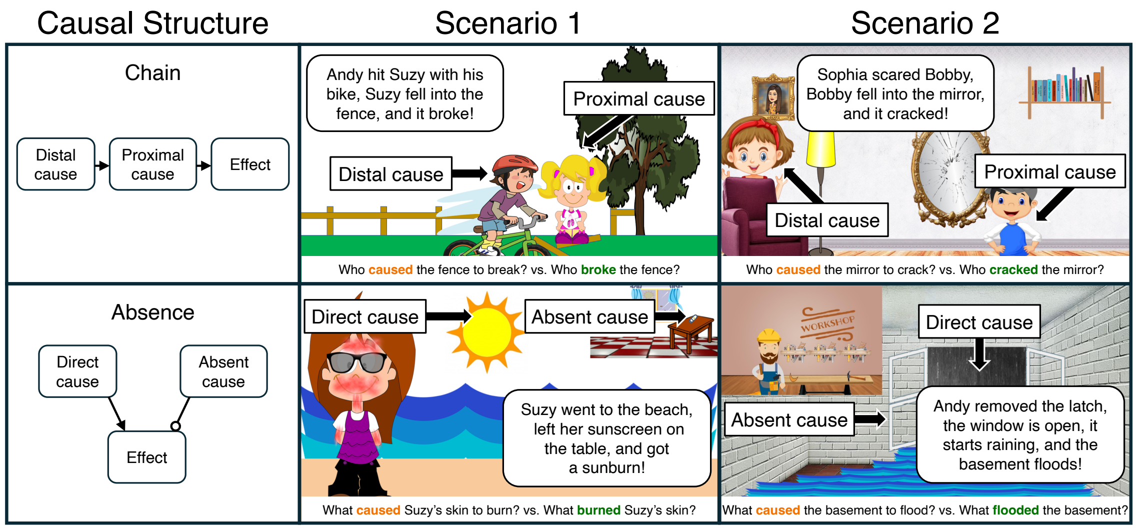

Figure 1

Experiment Overview: *Causal structures and illustrations of the final scene in different scenarios. The top row shows the chain cases from Experiment 1 and the bottom row shows the absence cases from Experiments 2 and 3. Participants were asked questions that either used a periphrastic causative (e.g., "caused to break") or a lexical causative (e.g., "broke"). In one condition, the periphrastic causative was "caused" (shown here), and in the other it was "made".*

are krading."). When asked to find "krading", children looked longer at the causal scene when hearing the transitive sentence and longer at the non-causal scene when hearing the intransitive sentence [see also 42, 43, 44, 45, 41, 46, 47]. When hearing a novel verb in a transitive sentence, children expect that verb to encode a causal relationship that involves spatiotemporal continuity [e.g., 48, 49, 50, 51, 52].

What about periphrastic causatives? Between the ages of 2 and 2½, children acquire one of their first periphrastic causatives, "made". Children then do two things. They *overlexicalize*, using a lexical causative when an adult would use a periphrastic construction (e.g., "Water bloomed these flowers" (= made these flowers bloom)) and *overanalyze*, using a periphrastic construction where a lexical causative is called for [e.g., "Then I'm going to sit on him and made him broken." (= break him); 53]. This suggests that young children have difficulty distinguishing between what kinds of causes lexical and periphrastic causatives refer to.

Semantics and pragmatics

While transitivity might help children distinguish statements that refer to causes from those that don't, it doesn't help with distinguishing what kinds of causes lexical and periphrastic causatives refer to. However, causatives differ in their specificity [54]. For example, "caused to break" is more general in that it's true of many possible ways in which breaking can happen, including directly and indirectly breaking something. In contrast, "break" is more specific—it picks out the subset of direct causes.

Because of the semantic overlap between lexical and periphrastic causatives, understanding causal verbs may be related to understanding scalar terms. For instance, "some" is less specific than "all". Though it's not literally false to say "some" even when "all" is true, using "some" pragmatically implies that "all" is not true (because the speaker could have said "all"). Consequently, adults don't find it acceptable for a speaker to say that "Some of the horses jumped over the fence." when in fact all of the horses did [see e.g., 55, for discussion].

In contrast, 5- to 6-year-old children say that a puppet "answered well" when all horses jumped over a fence and the puppet said "some of the horses jumped over the fence" [56, 55, 57]. Even though young children engage in various forms of pragmatic inference [58, 59, 60, 61], it seems that they understand "some" and "all" in ways that accord with their literal semantics. Across development, children learn the pragmatic implicature that's associated with using each expression, and they begin to understand them like adults.

Adults think it is less appropriate to say that "Suzy caused the fence to break" than "Suzy broke the fence", in the fence-breaking scenario [4]. But much like younger children think it's fine to say "some of the horses jumped over the fence" when all of the horses jumped over the fence, they might think the same for "Suzy caused the fence to break". Older children might realize that while this sentence is true, it's potentially misleading since a more specific and informative expression is available; namely that "Suzy broke the fence". Once children know that "caused to break" and "broke" can mean different things, they need to consider what pragmatic inferences a listener will make based on what they say.

Our question

As children develop, they first understand that causes directly produce effects, then that causes can be indirect, and finally that absences can also be causes [3]. How do children learn what causal expressions map onto these different kinds of causes?

How children map "caused" to different kinds of causes is of particular interest since we know so little about it. The word "caused" is almost never used in developmental work, even though it is commonplace in work on adult causal judgment [e.g., 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72]. Children are instead asked what "made" an outcome occur, whether they can "make" it happen [e.g., 73, 32, 35, 74, 75, 76, 77, 78], or whether something happened "because" of something else [e.g., 79, 80, 81, 82, 83]. It's understandable that "caused" doesn't feature prominently in developmental work. "Caused" is much less frequent than "made" in adult speech. According to the Corpus of Contemporary American English [84], "caused" is ranked 563 while "made" is ranked 50. Young children may not hear or produce the word "caused" very often.

Do children understand that "caused", "made", and lexical causatives can refer to different kinds of causes? There are two separate questions here. First, do children distinguish lexical and periphrastic causatives? Second, do children distinguish different periphrastic causatives? To draw different inferences about what lexical and periphrastic causatives refer to, children need to understand both the semantics of these verbs and the pragmatics of their use. For instance, in the fence-breaking scenario, children need to know that while it is true that both Andy and Suzy caused the fence to break, it is only true that Suzy broke it. If asked "Who caused the fence to break?", the better answer is Andy since the speaker could have used a more specific question—"Who broke the fence?"—if they wanted to hear about Suzy. Before children understand the semantics of these verbs and the pragmatics of their use, they might treat them as only referring to direct productive causes.

Children also need to understand that there are subtle semantic differences between periphrastic causatives, such as "caused" and "made" [see e.g. 85]. By examining both "caused" and "made" as periphrastic alternatives to lexical causatives, we ask whether children simply think any periphrastic causative refers to whatever the lexical causative does not, or whether they think that "made" and "caused" mean something different.

Results

To investigate how children develop their understanding of different causal expressions, we conducted three experiments examining their interpretation of lexical and periphrastic causatives across different causal scenarios. Before presenting the specific findings, we outline our statistical approach and preregistration procedures.

Statistical approach and preregistration

For all results reported in this paper, we analyzed the data using Bayesian logistic mixed effects models, using the default priors as specified by the `brms` package in `R` [86]. We will refer to a statistical result of interest as "credible" when the 95% credible interval excludes 0 (or 1 when reporting odds ratios).

Figure 1 and Table 1 show an overview of the experiments. We pre-registered separate analyses for each selected referent (e.g., "Andy"—distal cause; "Suzy"—proximal cause) and report these results in the Appendix. Here, we directly compare which referent was selected for a given causal verb, using the phrase "as predicted" to explicitly mark where our pre-registered hypotheses apply. All experiments, data, analyses, and links to pre-registrations are available here: `https://github.com/davdrose/cause_burn_development`

Experiment 1: Causal chains

The goal of this experiment was to determine whether adults and children understand "caused", "made" and lexical causatives ("broke", "cracked") to refer to different events in causal chains, such as the fence-breaking scenario (see Figure 1 top and Table 1). Who will children and adults refer to when asked "Who caused the fence to break?" and "Who broke the fence?"? And who will they refer to when asked "Who made the fence break?" and "Who broke the fence?"? Figure 2 shows the results.

"Caused" versus lexical causatives

As predicted, we found that children were more likely to select the distal cause (e.g., the person on the bike) when asked "Who caused the fence to break?" compared to "Who broke the fence?" (caused: 67%, 95% confidence interval (CI) [61%, 72%]; lexical: 20%, CI[16%, 25%], odds ratio: 13.8, 95% credible interval (CrI)[7.8, 22.4]). The same was true for adults (caused: 74%, CI[67%, 82%], lexical: 28%, CI[20%, 36%], odds ratio: 18.2, CrI [6.27, 40.7]). There was no credible effect of age on children's selections of the distal cause

Table 1

Overview of experiments. The scenario order was randomized in all experiments. In Experiments 1 and 2, the question order was also randomized. In Experiment 1, we counterbalanced which character appeared in which causal role (e.g., either Andy or Suzy was the one on the bike).

Experiments	Condition	Scenarios	Questions	Response Coding	Age (Children)	N (Children)	N (Adults)
Experiment 1	"Caused" versus lexical causatives ("break", "crack")	fence and mirror	"Who caused the fence to break?" and "Who broke the fence?" "Who caused the mirror to crack?" and "Who cracked the mirror?"	Andy (distal), Suzy (proximal) Sophia (distal), Bobby (proximal)	4 - 9	150	60
	"Made" versus lexical causatives ("break", "crack")	fence and mirror	"Who made the fence break?" and "Who broke the fence?" "Who made the mirror crack?" and "Who cracked the mirror?"	Andy (distal), Suzy (proximal) Sophia (distal), Bobby (proximal)	4 - 9	150	60
Experiment 2	"Caused" versus lexical causatives ("burn", "flood")	sunburn and flood	"What caused Suzy's skin to burn?" and "What burned Suzy's skin?" "What caused the basement to flood?" and "What flooded the basement?"	sun (direct), sunscreen (absent) rain/water (direct), latch/window (absent)	4 - 9	150	60
	"Made" versus lexical causatives ("burn", "flood")	sunburn and flood	"What made Suzy's skin burn?" and "What burned Suzy's skin?" "What made the basement flood?" and "What flooded the basement?"	sun (direct), sunscreen (absent) rain/water (direct), latch/window (absent)	4 - 9	150	60
Experiment 3	Absence with Explanation	sunburn and flood	"Why did Suzy get a sunburn?" "Why did the basement flood?"	sun (direct), sunscreen (absent) rain/water (direct), latch/window (absent)	4 - 6	61	30

(0.07, CrI [−.11, .26]), and there was no interaction between causal verb and age (0.21, CrI [−0.07, 0.48]).

"Made" versus lexical causatives

As predicted, children were more likely to select the distal cause for "made" compared to "lexical" (made: 47%, CI [41%, 52%], lexical: 22%, CI [17%, 27%], odds ratio: 4.8, CrI [2.8, 7.4]), and the same was true for adults (made: 52%, CI [43%, 61%], lexical: 28%, CI [19%, 36%], odds ratio: 4.6, CrI [2.0, 8.7]). There was no credible effect of age on selecting the distal cause (0.05, CrI [−0.14, 0.25]), but there was an interaction between causal verb and age (0.58, CrI [0.31, 0.86]) where selection of the distal cause decreased with age for "lexical" (−0.24, CrI [−0.49, 0.00]) but increased for "made" (0.34, CrI [0.12, 0.56]).

"Caused" versus "made"

We found that children were overall more likely to select the distal cause for "caused" compared to "made" (odds ratio: 2.61, CrI [1.76, 3.68]). The same was true for adults (odds ratio: 4.17, CrI [1.86, 7.63]).

Discussion

Like adults, 4-year-olds choose different referents in causal chain scenarios when a speaker uses "caused" versus a lexical causative. When Andy hits Suzy with his bike, she falls into the fence and it breaks, they think Andy, the distal cause, *caused* the fence to

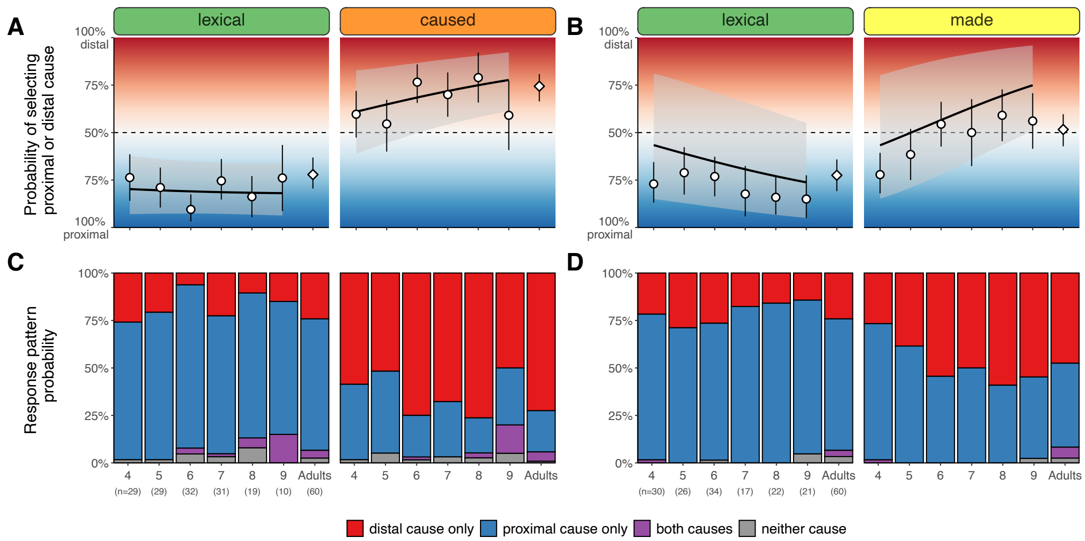

Figure 2

Experiment 1*: The top row (A and B) show the relative proportion with which participants selected the distal cause (red shading) versus the proximal cause (blue shading) depending on whether the speaker used a lexical causative or a periphrastic causative. For these we included all responses except for those where participants selected "neither" a distal nor a proximal cause. (A) shows the probability of selecting a proximal or distal cause in the "caused versus lexical" condition and (B) in the "made versus lexical" condition. Regression lines show the fits of Bayesian logistic mixed effects models with 80% credible intervals. Large points show the percentage with which each age group selected either referent. Error bars show 95% bootstrapped confidence intervals. The full response patterns are shown in the bottom row (C and D). (C) shows the response patterns in the "caused versus lexical" and (D) shows the "made versus lexical" condition for each age group along with the number of participants. Very few participants provided "neither cause" or "both causes" responses and most participants selected only a distal or proximal cause.*

break and that Suzy, the proximal cause, *broke* it. In contrast, "made" is treated more like a lexical causative early on. Children become less likely to think that "made" refers to proximal causes over development. But even adults don't think that "made" clearly refers to distal causes. In contrast, both children and adults think that "caused" refers to distal causes.

That children already understand "caused" to refer to causes that are more remote from their effects raises the possibility that they might even understand "caused" to refer to causes that are disconnected from their effects, like absences. We examine this in Experiment 2.

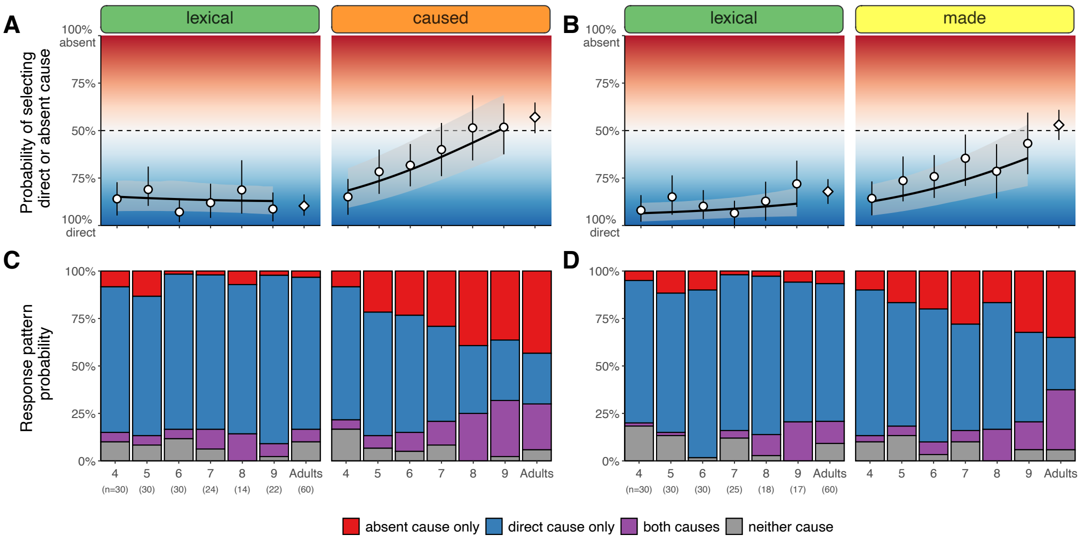

Figure 3

Experiment 2: *The top row (A and B) shows the relative proportion with which participants selected the absent (red shading) versus the direct cause (blue shading) depending on whether the speaker used a lexical causative or a periphrastic causative. For these we included all responses except for those where participants selected "neither" an absent nor a direct cause. (A) shows the probability of selecting a direct or absent cause in the "caused versus lexical" condition and (B) the "made versus lexical" condition. Regression lines show the fits of Bayesian logistic mixed effects models with 80% credible intervals. Large points show the percentage with which each age group selected either referent. Error bars show 95% bootstrapped confidence intervals. The full response patterns are shown in the bottom row (C and D). (C) shows the response patterns in the "caused versus lexical" and (D) shows the "made versus lexical" condition for each age group along with the number of participants. The majority of participants selected only an absent or direct cause.*

Experiment 2: Absences

This experiment investigates whether children understand "caused", "made" and lexical causatives to refer to different causes when presented with situations involving causation by absence, such as Suzy going to the beach, forgetting her sunscreen and getting a sunburn (see Figure 1 bottom and Table 1). What will children and adults answer when asked "What caused Suzy's skin to burn?" and "What burned Suzy's skin?"? And what will they answer when asked "What made Suzy's skin burn?" and "What burned Suzy's skin?"? Will they refer to an absence, such as the sunscreen, or a direct cause, such as the sun? Figure 3 shows the results.

"Caused" versus lexical causatives

As predicted, we found that children were more likely to select the absent cause when asked "What caused Suzy's skin to burn?" compared to "What burned Suzy's skin?"

(caused: 35%, CI [30%, 41%], lexical: 13%, CI [9%, 17%], odds ratio: 3.8, CrI [2.2, 5.6]). The same is true for adults (caused: 57%, CI [49%, 65%], lexical: 10%, CI [5%, 16%], odds ratio: 12.4, CrI [5.0, 23.2]). There was a credible effect of age on children's selections of the absent cause (.14, CrI [0, .31]), and there was an interaction between causal verb and age (.39, CrI [.08, .65]). Older children were more inclined to select the absent cause for "caused" (.33, CrI [.17, .49]) but not for "lexical" (−.05, CrI [−.29, .22]).

"Made" versus lexical causatives

In contrast to what we predicted, we found that children were more likely to select the absent cause for "made" compared to "lexical" (made: 27%, CI [22%, 32%], lexical: 12%, CI [8%, 16%], odds ratio: 3.4, CrI [1.9, 5.5]). The same is true for adults (made: 53%, CI [45%, 61%], lexical: 18%, CI [11%, 25%], odds ratio: 5.5, CrI [2.9, 9.4]). There was a credible effect of age on children's selections of the absent cause (.21, CrI [.02, .43]) but no interaction between causal verb and age [.15, CrI [−.15, .44]).

"Caused" versus "made"

We found that children were not more likely to select the absent cause for "caused" compared to "made" (odds ratio: 1.4, CrI [.93, 1.94]). The same was true for adults (odds ratio: 1.17, CrI [.67, 1.81]).

Discussion

When Suzy went to the beach, forgot her sunscreen and got a sunburn, young children almost never referred to the absence—the sunscreen—when asked what "caused" her skin to burn, "made" it burn, or "burned" it. Instead, they referred to the sun. With increasing age, children were more inclined to refer to an absence of sunscreen when they heard a periphrastic causative and were somewhat more likely to do so when the periphrastic causative was "caused" compared to "made".

While young children already have available a mapping of "caused" to distal causes in chains, the mapping of "caused" to absences shows protracted development. Why might this be the case? They might find it more challenging to draw a connection between the use of this word and this kind of cause. Or they might not even consider absences as causes. In Experiment 3, we ask whether children can conceive of absences as causes by simply asking them why something happened.

Experiment 3: Absences and explanation

Experiment 3 explores whether people are more likely to refer to absences when simply asked "why" an outcome occurred. Why-questions are often answered with "because", and adults are more likely to agree with sentences referring to absences when those sentences use "because" compared to "caused" [87]. Here, we test whether children might refer to absences when asked "why" an outcome occur. Figure 4 shows the results.

As predicted, when asked "Why did Suzy get a sunburn?", we found that children were more inclined to select absences, such as the sunscreen, than direct factors, such as the sun (absence: 73%, CI [66%, 79%], direct: 19%, CI [13%, 25%], odds ratio: 11.8, CrI [6.7, 18.7]. So were adults (absence: 93%, CI [87%, 100%], direct: 27%, CI [15%, 38%], odds

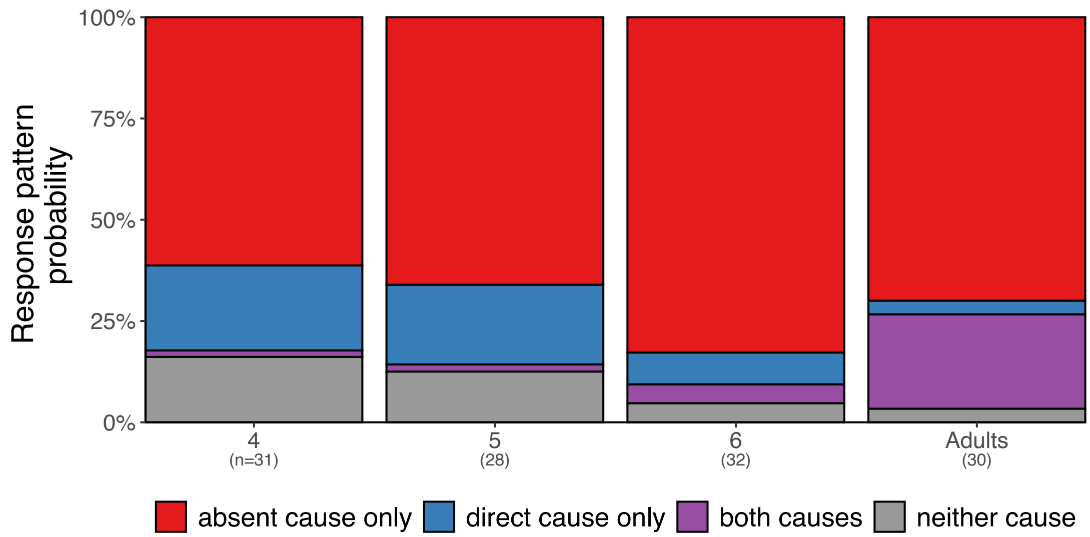

Figure 4

***Experiment 3**: Response patterns for each age group along with the number of participants. "both causes" means that they mentioned both the absent and direct cause, "neither cause" means they mentioned neither. Overall, adults and children tended to refer to an absence when asked "why" the outcome occurred. In contrast to children, adults were somewhat inclined to select "both causes", and children were somewhat more inclined than adults to select "neither cause".*

ratio: 53.5, CrI [9.1, 183]. Indeed, as the results show, even the youngest children in our sample were more likely to select absent than direct factors when asked a "why" question.

Discussion

When young children are told that Suzy goes to the beach, forgets her sunscreen and gets a sunburn and asked, "What caused Suzy's skin to burn", they tend to say "the sun". But when they are asked, "Why did Suzy's skin burn?", they say "because she forgot her sunscreen". This suggests that even though children can conceive of absences, they don't refer to absences when hearing the word "caused" until later.

General discussion

Linguists distinguish two kinds of causal expressions, lexical causatives and periphrastic causatives. Philosophers distinguish two kinds of causal relations, production and dependence. Recent work has found that adults use lexical causatives, like "burn", for productive causes that have a direct spatiotemporal connection to their effects and periphrastic causatives, like "cause", for dependence-based causes such as absences [4]. This suggests a deep connection between the different causal expressions that linguistics recognize and the different causal relations that philosophers have characterized. How do children learn to map different causal verbs to different kinds of causes?

Our findings reveal that young children understand that causal language picks out different causes. When considering a situation where Andy hits Suzy with his bike, Suzy falls into the fence and it breaks, 4-year-old children already refer to the proximal cause, Suzy (Experiment 1), when asked "Who broke the fence?". Likewise, when considering a situation where Suzy goes to the beach, forgets her sunscreen and gets a sunburn, they refer to the direct cause, the sun (Experiment 2), when asked "What burned Suzy's skin?". This suggests that young children treat lexical causatives, like "broke" and "burned", as picking out causal relations involving direct spatiotemporal contiguity.

Young children show an adult-like understanding of "caused" even though "caused" is rarer and less familiar than lexical causatives and other periphrastic causatives like "made". 4-year-old children already map "caused" to distal causes in a chain. When asked "Who caused the fence to break?", they refer to Andy, the person on the bike who hit Suzy (Experiment 1). So while children understand more frequent and familiar lexical causatives like "broke" to refer to proximal causes in chains, they also understand that "caused" can refer to distal causes in chains. At the same time they appreciate this, they don't distinguish "made" from lexical causatives, despite "made" being more frequent and familiar than "caused". Not even adults understand "made" to clearly refer to distal causes. Only "caused" does.

While young children map different causal verbs to different productive causes in chains, situations that involve absences are more challenging. When Suzy goes to the beach, forgets her sunscreen and gets a sunburn (Experiment 2), young children almost never refer to Suzy's having forgotten the sunscreen when asked "What caused her skin to burn?" or "What made her skin burn?". With increasing age, children become more inclined to do so and do so more readily for "caused" compared to "made". Yet children do think that absences can be causes before mapping "caused" to them. While 4-year-old children don't refer to an absence when asked, "What caused Suzy's skin to burn?", they do when asked "Why did Suzy's skin burn'?' (Experiment 3).

Together, our findings show that lexical causatives, like "broke" and "burned", and periphrastic causatives, like "made" and "caused" map to different causes over the course of development. What explains this developmental shift? We propose that both semantic and pragmatic mechanisms are at play.

Semantic development

Causal verbs differ in their specificity. Lexical causatives are more specific than "made" and "made" is more specific than "caused".

Lexical causatives refer to ways of directly causing particular outcomes. They specify many different direct, productive relations, like burning and breaking, that even young children understand. Across development, lexical causatives maintain a narrow scope—they only refer to direct productive causes, and not to indirect causes or absences. As such, they provide an early emerging, stable set of meanings that allow reference to many different and specific ways of producing outcomes.

Young children understand "made" much like they do "broke" or "burned": "made" refers to direct productive causes. Yet unlike lexical causatives, "made" undergoes some semantic expansion across development. Despite the fact that neither adults nor children tend to reliably refer to distal or absent causes when asked what "made" something happen, they are nonetheless more inclined to refer to these than when considering lexical causatives. But in contrast to the many different productive relations specified by lexical causatives, "made" provides a more general way to refer to productive causes, one that doesn't specify the particular manner in which the effect was produced. In this way, "made" might express a generic kind of production.

"Caused" has a more general meaning than "made"—one that goes beyond referring to productive causes. One possibility is that the meaning of "caused" is closely tied to that of counterfactual necessity, such that one event "caused" another to happen when it wouldn't have happened without it [see also e.g., 85]. This may be something children already recognize at 4-years-old when selecting distal causes. So when 4-year-old children refer to Andy as the one who "caused" the fence to break, they might appreciate that if Andy hadn't hit Suzy with his bike, the fence wouldn't have broken. Even if children have this meaning available for "caused" it may initially have a relatively narrow domain of application: it can refer to distal causes but not absences. So while 4-year-olds can think of absences—they cite Suzy forgetting her sunscreen when asked "why" she got a sunburn—they don't refer to them when asked what "caused" her skin to burn. They do so later, once the semantics of "caused" expands to include absences.

Pragmatic development

Semantic expansion is one key mechanism that can help explain how children learn to map causal verbs to different causes across development. Another mechanism is children's ability to draw pragmatic inferences about the use of different causal expressions in context. Here, we suggest a few factors that might change across development, using the Rational Speech Acts (RSA) framework as a guide [88, 89, 90].

Let's suppose for now that lexical causatives can only refer to direct causes, and that periphrastic causatives can refer to any kind of cause. Within the RSA framework, there are three factors that can help explain the patterns we observe: speaker optimality, utterance cost, and reference prior. Here, we outline the general idea, and provide a concrete implementation in the appendix. Figure 5A–C shows the model parameters (and how they are assumed to change across development), and Figure 5D shows how well this model captures the data in Experiments 1 and 2.

Speaker optimality captures how likely a speaker is to use more informative utterances. If older children assume that speakers are more likely to be informative, then this helps explain the general pattern that they are more likely to differentiate between the referents than younger children depending on whether they hear a lexical or periphrastic causative (Figure 5A).

The *reference prior* captures how likely a speaker is to refer to specific events. By assuming that speakers are more likely to refer to direct causes than distal causes, and even less likely to refer to absences, the model captures some of the differences between the scenarios (Figure 5B). Specifically, it accounts for the fact that when a periphrastic causative was used, participants were more likely to refer to distal causes in chains, than to absences in the other scenarios. Unlike the other factors in the model, the reference prior doesn't change with age.

Finally, the *utterance cost* captures how costly it is for a speaker to say something. An utterance is more costly when it's more difficult to say and when it comes to mind

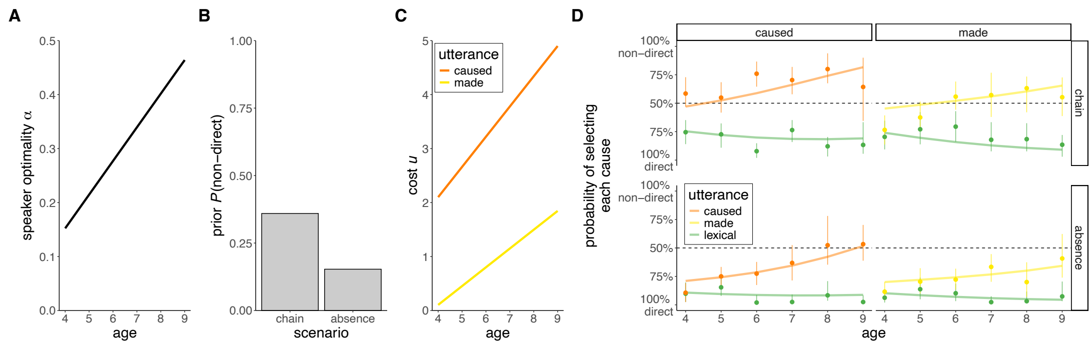

Figure 5

Rational Speech Act (RSA) model: A – C *Model parameters and how they change with development.* ***A*** *The speaker optimality parameter α increases with age, meaning that older children assume that speakers are more likely to produce more informative utterances.* ***B*** *The prior over the two referents is assumed not to change with age. The values show the prior probability of referring to the non-direct referent. This means that in the chain scenario, a speaker is more likely to refer to the "proximal cause" than the "distal cause". In the absence scenario, a speaker is even more likely to refer to the "direct cause" than the "absent cause".* ***C*** *The (relative) cost of using a periphrastic causative increases with age, and is greater for "caused" than for "made". The cost for producing the lexical causative is assumed to be 0.* ***D*** *Empirical data and model predictions. The points denote the proportion of participants who chose the direct cause or non-direct cause (i.e. the indirect or absent cause) as a referent depending on the utterance. The lines show the model predictions. On the left side, the possible utterances were "caused" vs. "lexical". On the right side, they were "made" vs. "lexical". The top row shows the "chain" scenarios, and the bottom row the "absence" scenarios. Error bars show 95% bootstrapped confidence intervals.*

less easily [88, 91]. Periphrastic causatives are longer expressions than lexical causatives (e.g., "caused to break" versus "broke"), so their utterance cost is higher. "Caused" is rarer than"made", so "caused" is more costly than "made". This difference in utterance cost between the expressions accounts for the fact that listeners are more likely to infer that the distal or absent cause was referred to when a speaker used "caused" compared to "made". If a speaker was willing to incur a greater cost, they must have wanted to be particularly informative (Figure 5C). The model assumes that, as children get older, they realize the greater relative cost of a periphrastic causative compared to a lexical causative. This could be because children need to learn which expressions are rare and which ones are more common. The model predictions and empirical data are shown in Figure 5D. The model accounts well for participants' responses, $r = 0.94, \text{RMSE} = 7.55$.

This model shows that even if the semantics of "caused" and "made" were identical, differences in their interpretation could arise from pragmatic effects like the ones we've outlined here. In reality, it is likely that both semantic and pragmatic factors jointly produce the changes in children's causal language understanding that we observed in our experiments. While the semantics of the model we consider here is fixed, the RSA framework can also model situations where the meaning of an expression needs to be learned, and might change over time [92]. In the appendix, we also look at how children's answers are affected by which causal expression they hear first. Order effects can provide further evidence for the role of pragmatic inference in language understanding [e.g., 93, 94, 95, 96].

Other verbs, other languages

Our experiments featured a small number of lexical causatives, and "caused" and "made" as periphrastic causatives. Of course, the English language features many more causatives. For example, in addition to "caused" and "made", English speakers can also say "enabled", "allowed", "let", etc. [see e.g., 97, 54]. It's possible that participants would be more inclined to refer to absences when asked what "allowed" it to happen rather than what "caused" it [97, 98]. If participants were instead asked, "What allowed Andy's basement to flood?" when he left the window open and it rained or if they were asked "What allowed Suzy's skin to burn?" when she went to the beach, forgot her sunscreen and got a sunburn, they may have been more inclined to refer to an absence—leaving the window open or forgetting the sunscreen—than when asked what "caused" these outcomes. Indeed, even though children show protracted development in mapping "caused" to absences, they have no trouble referring to them when asked "why" an outcome occurred, spontaneously using "because" to do so. It's possible that they map periphrastic causatives like "allowed" and "let" to absences earlier than they do so for "caused" or "made".

While we drew on the distinction between lexical and periphrastic causatives to explore how the connection between different causal verbs and relations arises, not all languages afford lexical and periphrastic constructions for expressing causation [99]. For instance, Mayan languages have morphological (e.g., a suffix can be added to a verb to give it a causal meaning) and periphrastic causatives, and Swahili has two kinds of morphological variants. Across languages, more direct kinds of causes are usually referred to by either lexical or morphological causatives. Indirect causes are typically referred to by periphrastic causatives or else by morphological causatives when the language uses lexical causatives for direct causes (e.g., Japanese). In cases where the language only has morphological causatives (e.g., Swahili) distinct types of morphological causatives are used for direct and indirect causes. While a number of languages use lexical or morphological causatives to refer to direct causes, an exception is Charu: periphrastic causatives refer to more direct causes. Indirect causes are referred to with morphological and periphrastic causatives [e.g., 15, 11]. Children who speak languages with different causal constructions from English, like Charu, might show different developmental patterns in mapping verbs to different causes [see e.g., 100, for work comparing German and Turkish speaking children)].

Conclusion

We examined how children develop a mapping from different causal verbs to different kinds of causes. They understand that lexical causatives, like "burned", refer to direct causes that produce their effects, that periphrastic causatives, like "caused", can to refer to causes that are more remote from their effects, and eventually come to understand that "caused" can also refer to absences. We argued that both semantic and pragmatic mechanisms contribute to this development. While more work is needed to tease these contributions apart, our findings suggest that semantic and pragmatic mechanisms support a distinction between causal verbs that involves mapping them to different causes. Our findings offer insight into the psychological mechanisms underlying the connection between the causal expressions that linguists recognize and the causal relations that philosophers characterize.

Methods

This research was approved by the Stanford Institutional Review Board (protocol no. IRB-59627), and all participants were asked to carefully read a consent form which included information about the study and the anonymity of the data. Adult participants read the form online and clicked a button indicating consent before proceeding to the experiment. Children were accompanied by a parent or guardian who verbally gave (recorded) consent and confirmed that the child assented to the study before it began. In addition, before the study began, children were given a description of the study, told that there were no right or wrong answers, and then decided to proceed to the study by clicking a button on the screen. Between test trials, children were also given a description of what was to come next, and decided to proceed by clicking a button on the screen. Children and parents (or guardians) were told that they could stop the study at any time, for any reason, and that there was no penalty for doing so. At the end of the study, parents or guardians also indicated their level of consent for how the video recording could be used (e.g., whether the video should be kept private or could be made public).

We preregistered sampling data from adults through Prolific [see 101, for details on demographics] and data from children through Lookit [see 102, for details on demographics] because these samples tend to be geographically and demographically diverse. For children, we preregistered that they would be included in the sample if they provided their own answers, without input from parents, siblings, etc., on all test trials.

In Experiments 1 and 2, children and adults were randomly assigned to the '"caused" versus lexical causatives' condition or the '"made" versus lexical causatives' condition. Each experiment included two scenarios and two questions that were presented in random order. In Experiment 3, children and adults read two scenarios and answered a single question. The scenarios were presented in a random order. In all experiments, we counterbalanced which character was in which causal role.

We collected data in Experiments 1 and 2 from 150 children for the '"caused" versus lexical causatives' condition and 150 children for the '"made" versus lexical causatives' condition with 30 in each age group from 4–6 and 60 in the 7–9 year old range. For Experiment 3, we preregistered that we would collect data from 60 children who were 4–6. For adults, we preregistered that we could collect data from 60 participants in the '"caused" versus lexical causatives' condition and in the '"made" versus lexical causatives' condition in both Experiments 1 and 2 (240 participants in total). In Experiment 3, we preregistered that we would collect data from 30 participants. For both children and adults, we recruited participants until the preregistered sample size was achieved. We didn't conduct power analyses to determine sample sizes. Our sample sizes were based on those of similar prior studies. For all experiments, data was collected from October 2021 to September 2023. Exclusions are reported in each study described below. All data was analyzed using `R` version 4.3.3.

Experiment 1: Causal chains

The "caused versus lexical" version for adult's was preregistered on October 28, 2022. The version for children was preregistered on November 1, 2022. The "made versus lexical" version for adults was preregistered on December 2, 2022. The version for children was preregistered on December 6, 2022.

Participants

We recruited 120 adult participants (Age: M = 35, SD = 13; Gender: 55 female, 56 male, 8 non-binary, 2 no response/other; Race: 12 Asian, 7 Black, 93 White, 9 no response/other) and 300 children (Gender: 128 female, 135 male, 37 no response/other). We pre-registered that we would collect data from 30 4-year-olds, 30 5-year-olds, 30 6-year-olds, and 60 7-9 year olds for each of the '"caused" versus lexical causatives' and the '"made" versus lexical causatives' condition. Our final sample includes only children who met our pre-registered inclusion criteria. For all experiments, each video was watched and it was confirmed whether the participants were children, and whether they completed it without interference from parents or siblings. 63 participants were excluded due to technical issues (no video, no audio) or failing to meet our pre-registered inclusion criteria (interference or incomplete). Adults were recruited through Prolific and paid at a rate of $12 an hour. Children were recruited through Lookit [102] and families were paid $5 for their participation.

Procedure

Children were tested asynchronously and began with warm-up trials, which were included to help them become comfortable with saying their answers out loud. After being introduced to Maggie, and being told that they would help her learn English, children were presented with two pairs of sentences: "I live in Maple Street/I live on Maple street" and "I put socks on my feet/I put socks on my feets". For each one, they were asked whether it is right or wrong for Maggie to say that. Children were asked to say their answers out loud.

Children then proceeded to the test scenarios (see Figure 1 top row and Table 1). In one, Andy hits Suzy with his bike, she falls into a fence and it breaks; in the other, Sophia hides behind a chair, jumps out, scares Bobby, he falls into a mirror and it cracks. In the '"caused" versus lexical causatives' condition, children were asked in the fence scenario "Who caused the fence to break?" and "Who broke the fence?". In the '"made" versus lexical causatives' condition, they were asked "Who made the fence break?" and "Who broke the fence?". In the mirror scenario, they were asked either "Who caused the mirror to crack?" (in the '"caused" versus lexical causatives' condition) or "Who made the mirror crack?" (in the '"made" versus lexical causatives' condition) as well as "Who cracked the mirror?". Children said their responses out loud.

The procedure for adults was the same except that they didn't complete warm-up trials, and instead of saying responses out loud, they wrote them in a text box. The adult version of the experiment was programmed in jsPsych [103].

Design

Participants were randomly assigned to the '"caused" versus lexical causatives' or '"made" versus lexical causatives' condition. In both conditions, the scenario and question order was randomized. We also counterbalanced which character was in which causal role (e.g., either Suzy or Andy was the one on the bike).

Response coding

We pre-registered coding responses into two categories: “distal” if the character on the bike (or behind the chair) was mentioned and “proximal” if the character who fell into the fence (or the mirror) was mentioned. Our coding scheme was not mutually exclusive. It was possible for a participant to refer to both a distal and proximal cause in the same response, or to neither of the two.

Experiment 2: Absences

The “caused versus lexical” version for adult’s was preregistered on October 6, 2021. The version for children was preregistered on March 7, 2022 (for 4-6 year olds) and on January 8, 2022 (for 7-9 year olds). The “made versus lexical” version for both adults as children was pre-registered on May 16, 2022.

Participants

We recruited 120 adult participants (Age: M = 30, SD = 10; Gender: 68 female, 49 male, 2 non-binary, 1 no response/other; Race: 12 Asian, 12 Black, 83 White, 13 no response/other) and 300 children (Gender: 147 female, 153 male). We pre-registered that we would collect data from 30 4-year-olds, 30 5-year-olds, 30 6-year- olds, and 60 7-9 year olds for each of the ‘“caused” versus lexical causatives’ and the ‘“made” versus lexical causatives’ conditions. Our final sample includes only children who met our preregistered inclusion criteria. 70 participants were excluded due to technical issues or failing to meet our pre-registered inclusion criteria. Adults were recruited through Prolific and paid at a rate of $12 an hour. Children were recruited through Lookit [102] and families were paid $5 for their participation.

Procedure

The procedure was the same as in Experiment 1. However, this time participants saw different scenarios (see Figure 1 bottom row and Table 1). In one of the scenarios, Suzy goes to the beach, forgets her sunscreen and gets a sunburn. In the other, Andy removes the latch from his basement window, the window is open when it starts to rain, and the basement floods. In the ‘“caused” versus lexical causatives’ condition, children were asked in the sunburn scenario “What caused Suzy’s skin to burn?” and “What burned Suzy’s skin?”. In the ‘“made” versus lexical causatives’ condition, they were asked “What made Suzy’s skin burn?” and “What burned Suzy’s skin?”. In the flood scenario, they were asked “What caused Andy’s basement to flood?” (in the ‘“caused” versus lexical causatives’ condition) and “What flooded Andy’s basement?” or “What made Andy’s basement flood?” (in the ‘“made” versus lexical causatives’ condition) and “What flooded Andy’s basement?” Children said their responses out loud.

Design

Participants were randomly assigned to the ‘“caused” versus lexical causatives’ or ‘“made” versus lexical causatives’ condition. In both conditions, the order of the scenarios and questions were randomized.

Response coding

We pre-registered coding responses into two categories: "absence" (e.g., if the sunscreen was mentioned) and "direct" (e.g., if the sun was mentioned). As in Experiment 1, our coding scheme was not mutually exclusive.

Experiment 3: Absences and Explanation

This experiment, for both adults and children, was preregistered on April 3, 2023.

Participants

We recruited 30 adult participants (Age: M = 34, SD = 12; Gender: 10 female, 16 male, 3 non-binary, 1 no response/other; Race: 6 Asian, 3 Black, 20 White, 1 no response/other) and 91 children (Gender: 53 female, 38 male). We pre-registered that we would collect data from 30 4-year-olds, 30 5-year-olds, and 30 6-year-olds. 18 participants were excluded due to technical issues or failing to meet our pre-registered inclusion criteria. Adults were recruited through Prolific and paid at a rate of $12 an hour. Children were recruited through Lookit [102] and families were paid $5 for their participation.

Procedure and design

The procedure was the same as in Experiment 2. However, this time participants were asked different questions. In the sunburn scenario, participants were asked, "Why did Suzy's skin burn?" and in the flood scenario, they were asked, "Why did the basement flood?". The order of the scenarios was randomized.

Response coding

We pre-registered the same coding scheme as in Experiment 2. As in our previous experiments, our coding scheme was not mutually exclusive.

Data availability

All data is available here: `https://github.com/davdrose/cause_burn_development`

Code availability

All code is available here: `https://github.com/davdrose/cause_burn_development`

Competing interests

The authors declare that there are no competing interests.

Acknowledgments

We would like to thank members of the Causality in Cognition Lab and the Markman Lab for their feedback, as well as Elizabeth Bonawitz, Eve Clark, Herb Clark, Mike Frank, Noah Goodman, Jonathan Kominsky, Beth Levin and Mahesh Srinivasan for helpful discussion. TG was supported by grants from Stanford's Human-Centered Artificial Intelligence institute (HAI) and Cooperative AI. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization: DR, SZ, SN, EM & TG; Methodology: DR, SZ, SN, EM & TG; Software: DR, SZ, TG; Validation: DR, TG; Formal Analysis: DR, TG; Investigation: DR, SZ; Data Curation: DR, SZ; Writing—Original Draft: DR; Writing—Review & Editing: DR, SN, EM, TG; Visualization: DR, TG; Supervision: DR, EM, TG; Project Administration; DR, EM, TG; Funding Acquisition: EM, TG.

The authors declare that there are no competing interests.

References

[1] David Hume. *A treatise of human nature*. Clarendon Press, Oxford, 1978 edition, 1789.

[2] Anne Schlottmann. Perception versus knowledge of cause and effect in children: When seeing is believing. *Current Directions in Psychological Science*, 10(4):111–115, 2001.

[3] Mariel K. Goddu and Alison Gopnik. The development of human causal learning and reasoning. *Nature Reviews Psychology*, pages 1–21, 2024.

[4] David Rose, Eric Sievers, and Shaun Nichols. Cause and burn. *Cognition*, 207:104517, 2021.

[5] Grace Song and Phillip Wolff. Linking perceptual properties to the linguistic expression of causation. In *Language, culture and mind*, pages 237–250. 2003.

[6] Jerry A. Fodor. Three reasons for not deriving "kill" from "cause to die". *Linguistic Inquiry*, 1(4):429–438, 1970.

[7] James D. McCawley. Conversational implicature and the lexicon. In Peter Cole, editor, *Syntax and semantics, vol. 9: Pragmatics*, pages 245–258. Academic Press, New York, 1978.

[8] Steven Pinker. *Learnability and cognition: the acquisition of argument structure*. MIT Press, Cambridge, MA, 1989.

[9] Masayoshi Shibatani. The grammar of causative constructions: a conspectus. In Masayoshi Shibatani, editor, *Syntax and semantics, vol. 6: The grammar of causative constructions*, pages 1–40. Academic Press, New York, 1976.

[10] Anna Wierzbicka. *The semantics of grammar*. John Benjamins, Amsterdam, 1988.

[11] I Nengah Aryawibawa, Yuli Qomariana, Ketut Artawa, and Ben Ambridge. Direct versus indirect causation as a semantic linguistic universal: Using a computational model of english, hebrew, hindi, japanese, and k'iche'mayan to predict grammaticality judgments in balinese. *Cognitive Science*, 45(4):e12974, 2021.

[12] Phillip Wolff. Direct causation in the linguistic coding and individuation of causal events. *Cognition*, 88(1):1–48, 2003.

[13] Anna Wierzbicka. Why “kill” does not mean “cause to die”: the semantics of action sentences. *Foundations of Language*, 13:491–528, 1975.

[14] Beth Levin and Malka Rappaport Hovav. Two structures for compositionally derived events. In *Semantics and linguistic theory*, pages 199–223, 1999.

[15] R. M. W. Dixon. A typology of causatives: Form, syntax and meaning. In R. M. W. Dixon and Alexandra Y. Aikhenvald, editors, *Changing valency: Case studies in transitivity*, pages 30–83. Cambridge University Press, New York, 2000.

[16] Charles J. Fillmore. Subjects, speakers, and roles. In Donald Davidson and Gilbert Harman, editors, *Semantics of natural language*, pages 1–24. Springer Netherlands, Dordrecht, 1972.

[17] N. Hall. Two concepts of causation. In J. Collins, N. Hall, and L. A. Paul, editors, *Causation and Counterfactuals*. MIT Press, 2004.

[18] Peter Godfrey-Smith. Causal pluralism. In H. Beebee, C. Hitchcock, and P. Menzies, editors, *Oxford Handbook of Causation*, pages 326–337. Oxford University Press, 2010.

[19] Phil Dowe. *Physical causation*. Cambridge University Press, Cambridge, 2000.

[20] Wesley Salmon. *Scientific explanation and the causal structure of the world*. Princeton University Press, Princeton, 1984.

[21] D. Lewis. Causation. *The Journal of Philosophy*, 70(17):556–567, 1973.

[22] Sara Bernstein. Omission impossible. *Philosophical Studies*, 173(10):2575–2589, 2016.

[23] David Lewis. Causation. *The journal of philosophy*, 70(17):556–567, 1973.

[24] Sarah McGrath. Causation by omission: A dilemma. *Philosophical Studies*, 123(1-2): 125–148, 2005.

[25] Jonathan Schaffer. Causation by disconnection. *Philosophy of Science*, 67(2):285, 2000. doi: 10.1086/392776. URL `http://dx.doi.org/10.1086/392776`.

[26] Paul Henne, Ángel Pinillos, and Felipe De Brigard. Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy*, 95(2):270–283, 2017.

[27] Tobias Gerstenberg and Stephan Stephan. A counterfactual simulation model of causation by omission. *Cognition*, 216:104842, 2021.

[28] Albert Michotte. *The Perception of Causality*. Routledge, 1963. doi: 10.4324/9781315519050.

[29] Alan M Leslie and Sarah Keeble. Do six-month-old infants perceive causality? *Cognition*, 25(3):265–288, 1987.

[30] Paul Muentener and Susan Carey. Infants' causal representations of state change events. *Cognition*, 61:63–86, 2010.

[31] Paul Muentener, Elizabeth Bonawitz, Alexandra Horowitz, and Laura Schulz. Mind the gap: Investigating toddlers' sensitivity to contact relations in predictive events. *PloS one*, 7(4):e34061, 2012.

[32] Elizabeth B Bonawitz, Daniel Ferranti, Rebecca Saxe, Alison Gopnik, Andrew N Meltzoff, James Woodward, and Laura E Schulz. Just do it? investigating the gap between prediction and action in toddlers' causal inferences. *Cognition*, 115(1):104–117, 2010.

[33] Merry Bullock and Rochel Gelman. Preschool children's assumptions about cause and effect: Temporal ordering. *Child Development*, pages 89–96, 1979.

[34] Teresa McCormack, Eimear O'Connor, Sarah Beck, and Aidan Feeney. The development of regret and relief about the outcomes of risky decisions. *Journal of Experimental Child Psychology*, 148:1–19, 2016. ISSN 00220965. doi: 10.1016/j.jecp.2016.02.008.

[35] Thomas R Shultz. Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47(1, Serial No. 194), 1982.

[36] Tim P German and Shaun Nichols. Children's counterfactual inferences about long and short causal chains. *Developmental Science*, 6(5):514–523, 2003.

[37] Michael Schleifer, Thomas R Shultz, and Monique Lefebvre-Pinard. Children's judgements of causality, responsibility and punishment in cases of harm due to omission. *British Journal of Developmental Psychology*, 1(1):87–97, 1983.

[38] Michael C Frank, Masha Braginsky, Daniel Yurovsky, and Virginia A Marchman. Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 2016.

[39] Melissa Bowerman. Learning the structure of causative verbs: A study in the relationship of cognitive, semantic, and syntactic development. *Papers and reports on child language development*, 8:142–178, 1974.

[40] György Gergely and Thomas G Bever. Related intuitions and the mental representation of causative verbs in adults and children. *Cognition*, 23(3):211–277, 1986.

[41] Letitia Naigles. Children use syntax to learn verb meanings. *Journal of child language*, 17(2):357–374, 1990.

[42] Ann Bunger and Jeffrey Lidz. Syntactic bootstrapping and the internal structure of causative events. In *Proceedings of the 28th annual boston university conference on language development*, volume 28, pages 74–85. Cascadilla Press Somerville, MA, 2004.

[43] Ann Bunger and Jeffrey Lidz. Constrained flexibility in the acquisition of causative verbs. In *Proceedings of the Annual Boston University Conference on Language Development*, volume 30, pages 60–71, 2006.

[44] Sudha Arunachalam, Emily Escovar, Melissa A Hansen, and Sandra R Waxman. Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and cognitive processes*, 28 (4):417–425, 2013.

[45] Sudha Arunachalam and Sandra R Waxman. Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114(3):442–446, 2010.

[46] Sylvia Yuan and Cynthia Fisher. "really? she blicked the baby?" two-year-olds learn combinatorial facts about verbs by listening. *Psychological science*, 20(5):619–626, 2009.

[47] Sylvia Yuan, Cynthia Fisher, and Jesse Snedeker. Counting the nouns: Simple structural cues to verb meaning. *Child development*, 83(4):1382–1399, 2012.

[48] Letitia R Naigles. The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, 58(2):221–251, 1996.

[49] Rose M Scott and Cynthia Fisher. Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and cognitive processes*, 24(6):777–803, 2009.

[50] Melissa Kline, Jesse Snedeker, and Laura Schulz. Linking language and events: Spatiotemporal cues drive children's expectations about the meanings of novel transitive verbs. *Language Learning and Development*, 13(1):1–23, 2017.

[51] Sudha Arunachalam and Shaun Dennis. Semantic detail in the developing verb lexicon: An extension of naigles and kako (1993). *Developmental science*, 22(1):e12697, 2019.

[52] Edward Kako. The semantics of syntactic frames. *Language and cognitive processes*, 21(5):562–575, 2006.

[53] Melissa Bowerman. Evaluating competing linguistic models with language acquisition data: Implications of developmental errors with causative verbs. *Quaderni di semantica*, 3:5–66, 1982.

[54] Ari Beller and Tobias Gerstenberg. Causation, meaning, and communication. *Psychological Review*, 2025.

[55] Ira A Noveck. When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2):165–188, 2001.

[56] Anna Papafragou and Julien Musolino. Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition*, 86(3):253–282, 2003.

[57] Yi Ting Huang and Jesse Snedeker. Semantic meaning and pragmatic interpretation in 5-year-olds: evidence from real-time spoken language comprehension. *Developmental psychology*, 45(6):1723, 2009.

[58] Alex J Stiller, Noah D Goodman, and Michael C Frank. Ad-hoc implicature in preschool children. *Language learning and development*, 11(2):176–190, 2015.

[59] Manuel Bohn, Michael H. Tessler, Michael Merrick, and Michael C. Frank. How young children integrate information sources to infer the meaning of words. *Nature Human Behaviour*, 5(8):1046–1054, 2021.

[60] Manuel Bohn, Michael H. Tessler, Michael Merrick, and Michael C. Frank. Predicting pragmatic cue integration in adults' and children's inferences about novel word meanings. *Journal of Experimental Psychology: General*, 151(11):2927, 2022.

[61] Manuel Bohn and Michael C Frank. The pervasive role of pragmatics in early language. *Annual Review of Developmental Psychology*, 1(1):223–249, 2019.

[62] C. R. Walsh and S. A. Sloman. The meaning of cause and prevent: The role of causal mechanism. *Mind & Language*, 26(1):21–52, 2011.

[63] Justin Sytsma, Jonathan Livengood, and David Rose. Two types of typicality: Rethinking the role of statistical typicality in ordinary causal attributions. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(4):814–820, 2011.

[64] Jonathan M Livengood, Justin Sytsma, and David Rose. Following the fad: Folk attributions and theories of actual causation. *Review of Philosophy and Psychology*, 8(2):273–294, 2017.

[65] M. D. Alicke. Culpable causation. *Journal of Personality and Social Psychology*, 63 (3):368–378, 1992.

[66] M. D. Alicke, D. Rose, and D. Bloom. Causation, norm violation, and culpable control. *The Journal of Philosophy*, 108(12):670–696, 2012.

[67] Jonathan F Kominsky and Jonathan Phillips. Immoral professors and malfunctioning tools: Counterfactual relevance accounts explain the effect of norm violations on causal selection. *Cognitive Science*, 43(11):e12792, 2019.

[68] C. Hitchcock and J. Knobe. Cause and norm. *Journal of Philosophy*, 11:587–612, 2009.

[69] Thomas F. Icard, Jonathan F. Kominsky, and Joshua Knobe. Normality and actual causal strength. *Cognition*, 161:80–93, 2017. doi: 10.1016/j.cognition.2017.01.010. URL `https://doi.org/10.1016%2Fj.cognition.2017.01.010`.

[70] Jana Samland and Michael R. Waldmann. How prescriptive norms influence causal inferences. *Cognition*, 156:164–176, 2016. doi: 10.1016/j.cognition.2016.07.007. URL `https://doi.org/10.1016%2Fj.cognition.2016.07.007`.

[71] Tania Lombrozo. Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, 61(4):303–332, 2010.

[72] D. A. Lagnado and S. Channon. Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, 108(3):754–770, 2008.

[73] Mary S. Ammon and Dan I. Slobin. A cross-linguistic study of the processing of causative sentences. *Cognition*, 7(1):3–17, 1979.

[74] Tamar Kushnir and Alison Gopnik. Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology*, 43(1):186, 2007.

[75] A. Gopnik, D.M. Sobel, L.E. Schulz, and C. Glymour. Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, 37(5):620–629, 2001.

[76] David M. Sobel, Joshua B. Tenenbaum, and Alison Gopnik. Children's causal inferences from indirect evidence: Backwards blocking and bayesian reasoning in preschoolers. *Cognitive science*, 28(3):303–333, 2004.

[77] Laura E. Schulz and Jessica Sommerville. God does not play dice: Causal determinism and children's inferences about unobserved causes. *Child Development*, 77(2):427–442, 2006.

[78] Jana Samland, Marina Josephs, Michael R. Waldmann, and Hannes Rakoczy. The role of prescriptive norms and knowledge in children's and adults' causal selection. *Journal of Experimental Psychology: General*, 145(2):125–130, 2016. doi: 10.1037/xge0000138. URL `https://doi.org/10.1037%2Fxge0000138`.

[79] Lois Hood and Lois Bloom. What, when, and how about why: A longitudinal study of early expressions of causality. *Monographs of the Society for Research in Child Development*, 44(6):1, 1979. doi: 10.2307/1165989. URL `http://dx.doi.org/10.2307/1165989`.

[80] Linda J Bebout, Sidney J Segalowitz, and George J White. Children's comprehension of causal constructions with "because" and "so". *Child Development*, pages 565–568, 1980.

[81] Allyssa McCabe and Carole Peterson. A naturalistic study of the production of causal connectives by children. *Journal of Child Language*, 12(1):145–159, 1985.

[82] Carole Peterson and Allyssa McCabe. Understanding "because": How important is the task? *Journal of Psycholinguistic Research*, 14:199–218, 1985.

[83] Anne K Hickling and Henry M Wellman. The emergence of children's causal explanations and theories: evidence from everyday conversation. *Developmental psychology*, 37(5):668, 2001.

[84] Mark Davies. The corpus of contemporary american english (coca), 2008. Available online at https://www.english-corpora.org/coca/.

[85] Prerna Nadathur and Sven Lauer. Causal necessity, causal sufficiency, and the implications of causative verbs. *Glossa: a journal of general linguistics*, 5(1), 2020.

[86] Paul-Christian Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.

[87] Jonathan Livengood and Edouard Machery. The folk probably don't think what you think they think: Experiments on causation by absence. *Midwest Studies in Philosophy*, 31(1):107–127, 2007.

[88] Judith Degen. The rational speech act framework. *Annual Review of Linguistics*, 9 (1):519–540, 2023.

[89] Noah D. Goodman and Michael C. Frank. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11):818–829, 2016. doi: 10.1016/j.tics.2016.08.005. URL `https://doi.org/10.1016%2Fj.tics.2016.08.005`.

[90] Michael C Frank and Noah D Goodman. Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998, 2012.

[91] Judith Degen, Robert D. Hawkins, Caroline Graf, Elisa Kreiss, and Noah D. Goodman. When redundancy is useful: A bayesian approach to "overinformative" referring expressions. *Psychological Review*, 127(4):591–621, jul 2020. doi: 10.1037/rev0000186. URL `https://doi.org/10.1037%2Frev0000186`.

[92] Leon Bergen, Roger Levy, and Noah Goodman. Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, 9:20–1, 2016.

[93] Eve V Clark. On the pragmatics of contrast. *Journal of child language*, 17(2):417–431, 1990.

[94] Roberta Michnick Golinkoff, Roberta Church Jacquet, Kathy Hirsh-Pasek, and Ratna Nandakumar. Lexical principles may underlie the learning of verbs. *Child development*, 67(6):3101–3119, 1996.

[95] Gil Diesendruck and Lori Markson. Children's avoidance of lexical overlap: a pragmatic account. *Developmental psychology*, 37(5):630, 2001.

[96] Justin Halberda. The development of a word-learning strategy. *Cognition*, 87(1): B23–B34, 2003.

[97] P. Wolff, A. K. Barbey, and M. Hausknecht. For want of a nail: How absences cause events. *Journal of Experimental Psychology: General*, 139(2):191–221, 2010.

[98] Huseina Thanawala and Christopher D Erb. Revisiting causal pluralism: Intention, process, and dependency in cases of double prevention. *Cognition*, 248:105786, 2024.

[99] Masayoshi Shibatani and Prashant Pardeshi. The causative continuum. In Masayoshi Shibatani, editor, *The grammar of causation and interpersonal manipulation*, pages 85–126. John Benjamins, Amsterdam, 2002.

[100] Ebru Ger, Larissa Stuber, Aylin C Küntay, Tilbe Göksun, Sabine Stoll, and Moritz M Daum. Influence of causal language on causal understanding: A comparison between swiss german and turkish. *Journal of Experimental Child Psychology*, 210:105182, 2021.

[101] Benjamin D. Douglas, Patrick J. Ewell, and Markus Brauer. Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *Plos one*, 18(3):e0279720, 2023.

[102] Kimberly Scott and Laura Schulz. Lookit (part 1): A new online platform for developmental research. *Open Mind*, 1(1):4–14, 2017.

[103] J R de Leeuw. jspsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47(1):1–12, 2015.

Broke or caused to break: How children map causal verbs to different causes across development

David Rose[1], Siying Zhang[2], Shaun Nichols[3], Ellen M. Markman[1], and Tobias Gerstenberg[1]
[1]Department of Psychology. Stanford University
[2]Department of Psychology, University of Washington
[3]Department of Philosophy, Cornell University

Contents

Experiment 1: Chains . . . 2
“Caused” versus lexical causatives . . . 2
“Made” versus lexical causatives . . . 3
“Caused” versus “made” . . . 3
Experiment 2: Absences . . . 3
“Caused” versus lexical causatives . . . 3
“Made” versus lexical causatives . . . 4
“Caused” versus “made” . . . 5
Order Effects . . . 5
Experiment 1: Chains . . . 5
Experiment 2: Absences . . . 6
A rational speech acts model of how children understand causal expressions . . . 7
Literal listener . . . 8
Speaker . . . 9
Pragmatic listener . . . 9
Fitting the model to our data . . . 10

References **12**

Appendix A
Preregistered Analyses

We preregistered separate hypotheses and analyses for each kind of cause in Experiment 1 and Experiment 2. We report the results here.

We fit Bayesian logistic mixed effects models and will refer to a statistical result of interest as "credible" when the 95% credible interval excludes 0 (except when reporting odds ratios, which we will interpret as credible when the interval excludes 1).

Experiment 1: Chains

Figure A1 shows the results.

"Caused" versus lexical causatives

We predicted that, for distal causes, there would be an interaction between causal verb and age, with children becoming more inclined to select distal causes when the verb was "caused". But we didn't find an interaction (.12, CrI [−.01, .27]). Even the youngest children in our sample—4-year-old children—were more inclined to select distal causes when asked, for example, "Who caused the fence to break?" as opposed to, for example, "Who broke the fence?". We also predicted that, for proximal causes, there would be an interaction between causal verb and age but didn't find credible evidence of this (−.15, CrI [−.31, .00]).

Children were, as we predicted, overall more inclined to select distal causes, such as the person on the bike, when "caused" was used (1.28, CrI [1.04, 1.53]), and more inclined to select proximal causes when the question used a lexical causative (−1.42, CrI [−1.71, −1.15]). Moreover, children's tendency to select a distal (.11, CrI [−.07, .29]) or proximal cause (−.02, CrI [−.23, .19]) didn't increase with age, which was in line with our prediction for proximal, but not distal causes.

In separately analyzing the data for adults, we found, as predicted, that they were also more inclined to understand "caused" to refer to distal causes (1.63, CrI [1.18, 2.14])

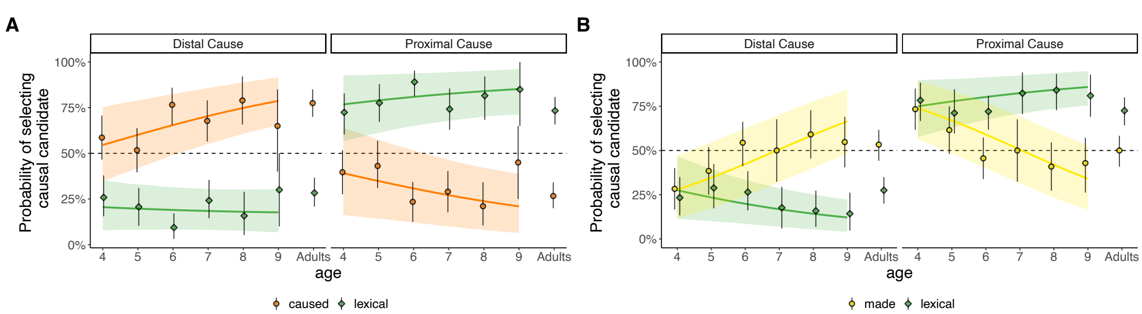

Figure A1

Experiment 1: *Probability of selecting a distal or proximal cause in the '"caused" versus lexical causatives' condition (A) and the '"made" versus lexical causatives' condition (B). Regression lines show the fits of Bayesian logistic mixed effects models with 80% credible intervals. Large points show the percentage with which each age group selected either referent. Error bars show 95% bootstrapped confidence intervals.*

and lexical causatives to refer to proximal causes (−1.73, CrI [−2.32, −1.23]).

“Made” versus lexical causatives

As predicted, there was an interaction between age and causal verb when contrasting “made” and lexical causatives for both the distal (.29, CrI [.16, .43]) and proximal (−.26, CrI [−.40, −.13]) cause. In contrast to “caused”, the youngest children did not distinguish “made” from lexical causatives like “broke”. For both questions they referred to the person who fell into the fence.

Children were, as predicted, overall more inclined to understand “made” to refer to distal causes (.79, CrI [.57, 1.02]) and lexical causatives to refer to proximal causes (−.75, CrI [−.98, -.53]). But their tendency to select the distal (.03, CrI [−.16, .22]) or proximal cause (−.09, CrI [−.29, .09]) didn’t increase with age, which was in line with our prediction for the proximal but not the distal cause.

Adults were, as predicted, more inclined to select a distal cause for “made” (.91, CrI [.54, 1.31]) and a proximal cause when the question used a lexical causative (−.73, CrI [−1.10, −.39]).

“Caused” versus “made”

In line with our predictions, children were more inclined to select distal causes for “caused” compared to “made” (1.14, CrI [.53, 1.83]) and more inclined to select proximal causes for “made” compared to “caused” (−1.40, CrI [−2.17, −.72]). We also found, as predicted, that children were more inclined to select distal causes with increasing age (.39, CrI [.14, .65]) and that, in contrast to our prediction, children became less likely to select proximal causes with increasing age (−.44, CrI [−.74, −.15]). Also, we predicted that there would be an interaction between the periphrastic causative and age but didn’t find evidence of this (distal: −.09, CrI [−.47, .27], proximal: .21, CrI [−.22, .64]).

For adults, they were, as predicted, more inclined to select distal causes for “caused” (2.23, CrI [.90, 3.91]). And though we predicted that there would be no credible difference in the extent to which adults selected proximal causes based on the verb used in the periphrastic construction, we found that they were actually more likely to select proximal causes for “made” (−1.78, CrI [−3.10, −.72]).

Experiment 2: Absences

Figure A2 shows the results.

“Caused” versus lexical causatives

We predicted that, for absent causes, there would be an interaction between causal verb and age, with children becoming more inclined to select absent causes when the verb was “caused”. This is what we found (.32, CrI [.17, .47]). The youngest children in our sample did not distinguish “caused” from lexical causatives like “burned” and almost never selected absences as causes. However, with age, children became more inclined to select an absence, such as the lack of sunscreen, for “caused” than when asked a question using a lexical causative. We also predicted that, for direct causes, there would not be an interaction between causal verb and age. We found, however, that there was an interaction (−.24,

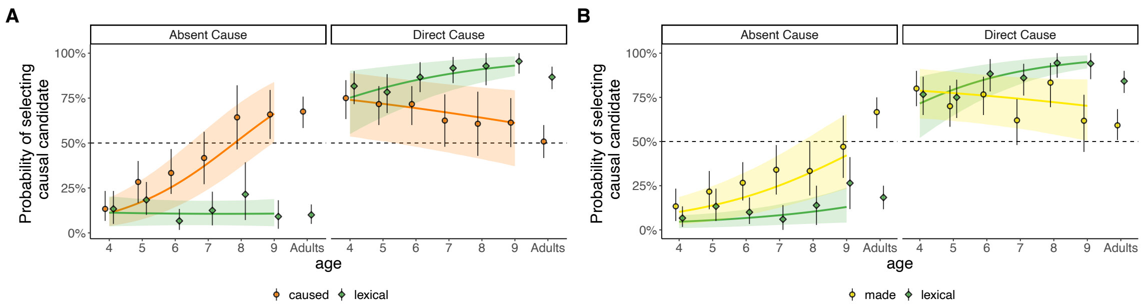

Figure A2

Experiment 2: *Probability of selecting an absent or direct cause in the '"caused" versus lexical causatives' condition (A) and the '"made" versus lexical causatives' condition (B). Regression lines show the fits of Bayesian logistic mixed effects models with 80% credible intervals. Large points show the percentage with which each age group selected either referent. Error bars show 95% bootstrapped confidence intervals.*

CrI [−.40, −.10]), with children being more inclined, as they got older, to select direct causes, such as the sun, for the question using the lexical causative.

Children were, as we predicted, overall more inclined to select absent causes when the question used "caused" (.86, CrI [.61, 1.12]), and to select direct causes when the question used a lexical causative (−.71, CrI [−.96, −.47]). Moreover, children's tendency to select the absent causes (.10, CrI [.10, .52]) increased with age but the tendency to select direct causes did not (.11, CrI [−.08, .30]). This was in line with our prediction for absent but not direct causes.

In separately analyzing the data for adults, we found, as predicted, that they were also more inclined to understand "caused" to refer to absent causes (1.80, CrI [1.34, 2.35]) and lexical causatives to refer to direct causes (−.98, CrI [−1.33, −.64]).

"Made" versus lexical causatives

As predicted, there was no interaction between age and causal verb when contrasting "made" and lexical causatives for the absent cause (.08, CrI [−.07, .24]). Children didn't credibly distinguish "made" from lexical causatives with increasing age in the extent to which they selected absences as causes. We predicted that there wouldn't be an interaction between causal verb and age for direct causes but we instead found evidence of an interaction (−.29, CrI [−.45, −.13]), with children being more inclined to select direct causes when asked a question using the lexical causative.

Children were overall more inclined to understand "made" to refer to absent causes (.73, CrI [.45, 1.02]) and lexical causatives to refer to direct causes (−.54, CrI [−.79, −.29]) but we predicted that there would be no difference for either cause. Their tendency to select the absent (.35, CrI [.11, .61])—but not direct causes (.18, CrI [−.03, .40])—increased with age, which was in line with our prediction for the absent but not the direct cause.

In contrast to our predictions, adults were more inclined to appeal to absences when asked a question using "made" (1.69, CrI [1.20, 2.25]) and they were more inclined to select direct causes when asked the question that only featured the lexical causative (−.80, CrI [−1.16, −.45]).

"Caused" versus "made"

Children were not, in contrast to what we predicted, more inclined to select absent causes when the periphrastic construction used "caused" compared to "made" (.65, CrI [−.01, 1.33]) and they didn't, as we predicted, differ in the extent to which they selected direct causes depending on the verb used (−.25, CrI [−.75, .24]). We also found, as predicted, that children were more inclined to select absent causes with increasing age (.45, CrI [.18, .76]) and that, in contrast to our prediction, children didn't become more likely to select direct causes with increasing age (−.10, CrI [−.31, .11]). Also, we predicted that there would be an interaction between the periphrastic causative and age but didn't find evidence of this [absent: .29, CrI [−.10, .70], direct: −.04, CrI [−.34, .25]).

For adults, we predicted that they would be more inclined to select absences for "caused" but we found instead that there was no credible difference between "caused" and "made" (.05, CrI [−.86, .98]). We found evidence, as predicted, that they didn't distinguish "caused" and "made" for direct causes (−.40, CrI [−1.08, .24]).

Appendix B
Pragmatics

We explore pragmatics in two different ways. First, we examine order effects. Then we describe a rational speech act model of pragmatic development.

Order Effects

We conducted exploratory analyses of order effects as a way to examine the influence of pragmatic factors on participants' responses. Taking causal chains as an example, like Andy hitting Suzy with his bike, Suzy falling into the fence, and the fence breaking, suppose that "caused" can refer to either the distal or proximal cause. And suppose that lexical causatives, like "broke", can only refer to the proximal cause. If someone first asked who "caused" the fence to break, then it would be unclear who they are referring to. If instead they asked first who "broke" the fence, it would be clear that they referred to the proximal cause. If they asked next, who "caused" the fence to break, then a pragmatic inference would suggest that they intended to refer to the distal cause this time. So the order of the questions might affect how reference is resolved when determining who "caused" (or "made") an outcome occur. Accordingly, we would expect more uncertainty about the referent when a periphrastic causative comes first as compared to when it comes second. We explore whether there were such order effects in Experiment 1 and Experiment 2.

Experiment 1: Chains

Due to a coding error in the adult version of the experiment, we are unable to recover the question order. We thus restrict our attention to children, examining order effects for responses to the first scenario that a participant saw.

The results are shown in Figure B1. For the '"caused" versus lexical causatives' condition, we found that children were not more inclined to think that "caused" referred

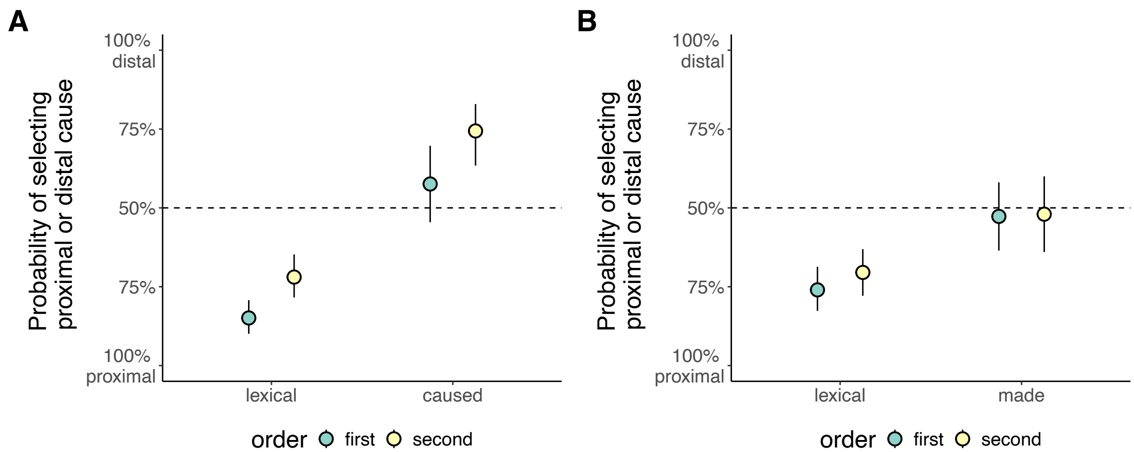

Figure B1

***Experiment 1**: Probability of selecting a proximal or distal cause in the '"caused" versus lexical causatives' condition (A) and the '"made" versus lexical causatives' condition (B), depending on whether a question was asked first or second. Points show the percentage with which children selected either referent. Error bars show 95% bootstrapped confidence intervals.*

to the distal cause when it was asked second (first: 56%, confidence interval (CI) [45%, 70%], second: 74%, confidence interval (CI) [65%, 84%], odds ratio: 1.96, credible interval (CrI) [.7, 3.96]). For the lexical causative, they were also not more inclined to think that it referred to the distal cause when it was asked second (first: 17%, CI [8%, 25%], second: 30%, CI [19%, 41%], odds ratio: 2.89, CrI [.85, 6.32]).

For the '"made" versus lexical causatives' condition, we again found that children were not more likely to think "made" referred to distal causes when it was asked second (first: 47%, CI [36%, 59%], second: 48%, CI [36%, 60%], odds ratio: 1.19, CrI [.29, 2.71]). For the lexical causative, they were also not more inclined to think that it referred to the distal cause when it was asked second (first: 24%, CI [14%, 34%], second: 30%, CI [19%, 40%], odds ratio: 1.21, CrI [.23, 3.19]).

Overall, there was no evidence that question order had an effect on responses.

Experiment 2: Absences

The results are shown in Figure B2. In the '"caused" versus lexical causatives' condition, children were not more inclined to think that "caused" referred to the absent cause when it came second (first: 51%, CI [39%, 62%], second: 24%, CI [15%, 33%], odds ratio: 3.46, CrI [1.45, 6.67]). They also didn't think the lexical causative was more likely to refer to an absence whether it was asked first or second (first: 13%, CI [5%, 20%], second: 10%, CI [3%, 18%], odds ratio: 1.25, CrI [.26, 3.13]). Adults didn't display any order effects (caused: first 48%, CI [29%, 68%], second 58%, CI [41%, 74%], odds ratio: 1.49, CrI [.31, 3.74]; lexical: first 6%, CI [0%, 14%], second 14%, CI [0%, 29%], odds ratio: 2.83, CrI [0.3, 17.78]).

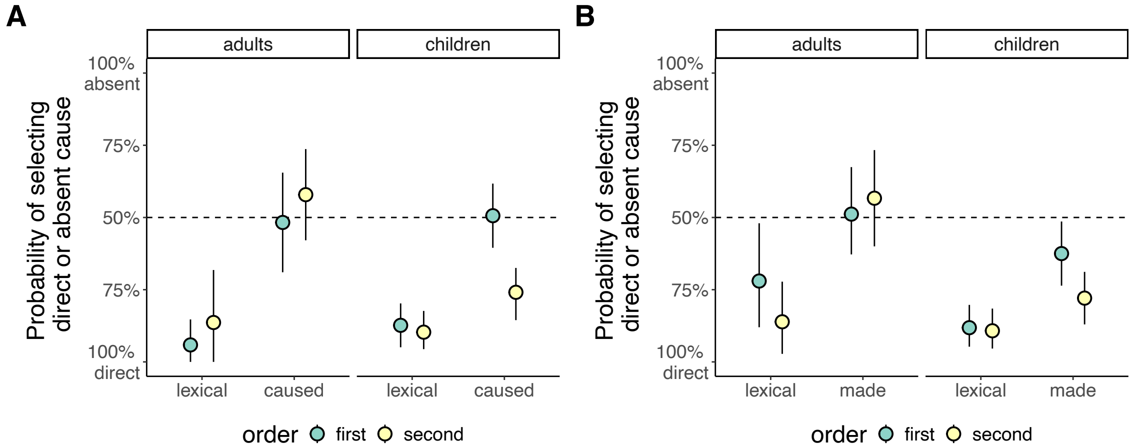

Figure B2

Experiment 2: *Probability of selecting a direct or absent cause in the '"caused" versus lexical causatives' condition (A) and '"made" versus lexical causatives' condition (B), depending on whether a question was asked first or second. Points show the percentage with which each age group selected either referent. Error bars show 95% bootstrapped confidence intervals.*

In the '"made" versus lexical causatives' condition, children were not more inclined to think that "made" referred to the absent cause when it came second (first: 38%, CI [26%, 49%], second: 22%, CI [13%, 32%], odds ratio: 2.67, CrI [.82, 6.59]). They also were not more inclined to think that the lexical causative referred to the absent cause when it came second (first: 12%, CI [4%, 19%], second: 11%, CI [3%, 19%], odds ratio: 1.10, CrI [.15, 3.40]). Adults didn't display any order effects (made first: 51%, CI [36%, 67%], made second: 57%, CI [37%, 79%], odds ratio: 1.05, CrI [.24, 2.55]; lexical first: 28%, CI [9%, 47%], lexical second: 14%, CI [2%, 26%], odds ratio: .42, CrI [.04, 1.36]).

In contrast to chains, we did find evidence that order had an effect on responses to absences. In particular order had an effect on responses to "caused". But it was opposite of what we would have expected, given that children were more inclined to think that "caused" referred to an absence when they were asked the "caused" question first.

A rational speech acts model of how children understand causal expressions

We model the development of children's understanding of causal expressions using the rational speech acts (RSA) framework (Degen, 2023; Frank & Goodman, 2012; Goodman & Frank, 2016). This framework formalizes the maxims of Gricean communication (Grice, 1975), such as that a speaker is likely to choose utterances that are truthful, and informative to a listener. The listener then interprets the meaning of these utterances in context. Specifically, we model our setting as a reference game where the speaker sends a message by choosing an utterance, and the listener tries to interpret what causal event the speaker referred to. Pragmatic inference emerges in the RSA via a process of recursive reasoning (e.g., Yoshida, Dolan, & Friston, 2008). We model participants as pragmatic listeners who reason about a speaker, who in turn reasons about a literal listener. We will describe how pragmatic reasoning emerges in that chain of reasoning, starting with the literal listener.

Literal listener

The literal listener takes in an utterance u and maps it onto a set of referents r. In our case, the possible set of utterances is either {"caused", "lexical"}, or {"made", "lexical"}, depending on the experimental condition. While the type of lexical utterance differs between each scenario (e.g., it's "burned" in one and "cracked" in the other), we model them under the same header here. The possible referents are the different candidate causes. In the causal chain scenarios, possible referents are {"proximal cause", "distal cause"}. In the absence scenarios, the possible referents are {"direct cause", "absent cause"}. Here, we treat the "proximal cause" and "direct cause" the same and call it the "direct cause". Analogously, we treat the "distal cause" and the "absent cause" the same, and call it the "non-direct cause".

Table B1

Semantics of the different causal utterances. "caused" and "made" can refer to both direct and non-direct causes. "lexical" expressions can only refer to direct causes.

	utterances u		
referents r	caused	made	lexical
direct	1	1	1
non-direct	1	1	0

This means that the speaker has two possible utterances, either "caused" versus "lexical", or "made" versus "lexical" (depending on the experimental condition), and there are two possible referents, either the "direct cause", or the "non-direct' cause. Table B1 shows the literal meaning of the different causal utterances. Accordingly, the periphrastic causatives "caused" and "made" are true for both direct and non-direct causes. However, lexical causatives (such as "broke", or "burned") are only true for direct causes. Intuitively, this means that when a speaker uses a lexical causative, it's clear what she meant: she refers to the direct cause. In contrast, when she uses a periphrastic causative, there is some ambiguity about what she meant as the literal meaning is consistent with both the direct and the non-direct cause.

Formally, the literal listener L_0 is defined as follows:

$$P_{L_0}(r \mid u) = \delta_{r \in [\![u]\!]} \cdot P(r). \tag{B1}$$

He takes in an utterance u and infers the referent r by considering the literal semantics (see Table B1) as well as the prior probability with which a speaker is likely to refer to the different referents $P(r)$. The literal semantics is captured in the meaning function $\delta_{r \in [\![u]\!]}$ which just checks what meaning is true for a given utterance.

So, assuming that the prior is $P(r) = 0.5$, which means that the referent is just as likely to be the direct cause or the non-direct cause a priori, when a literal listener hears "caused" or "made", they would be equally likely to think that it refers to each of the two possible causes.

Speaker

A speaker S wants to communicate a referent to the listener and decides what utterance to choose based on its expected utility. In principle, many considerations can go into this utility. Here, we consider that a speaker chooses an utterance so as to be informative and to minimize her cost of producing the utterance. Formally, the utility $U(u, r)$ is defined as

$$U(u, r) = \ln P_{L_0}(r \mid u) - \text{cost}(u), \tag{B2}$$

where $P_{L_0}(r \mid u)$ captures what inference the literal listener would make based on the utterance, and $\text{cost}(u)$ is the cost of that utterance. There are several ways to think of what the utterance cost means. Generally, shorter utterances are less costly than longer utterances. If a speaker can convey the same information saying less, then she should do so. In general, periphrastic causatives (e.g., "caused to crack") are longer than lexical causatives (e.g., "cracked"). Moreover, it's also often assumed that word frequency affects utterance cost. Expressions that are rare have a greater utterance cost than expressions that are common (e.g., because more frequent expressions come to mind more easily, see Degen, 2023; Degen, Hawkins, Graf, Kreiss, & Goodman, 2020).

The speaker chooses her utterance in a soft-max optimal way (Sutton & Barto, 1998), like so

$$P_S(u \mid r) = \frac{\exp(\alpha \cdot U(u, r))}{\sum_{i=1}^{n} \exp(\alpha \cdot U(u_i, r))}. \tag{B3}$$

The α parameter captures the speaker optimality. A high α value means that the speaker is very likely to use the utterance with the greater utility. If $\alpha = 0$, she would chose each utterance with the same probability, even when their utilities differ. The denominator sums over the possible utterances u_i that the speaker could use.

Pragmatic listener

Finally, the pragmatic listener P_{L_1} considers the utterance u that the speaker S produced and makes an inference about the intended referent r using Bayes' rule. Formally,

$$P_{L_1}(r \mid u) = \frac{P_S(u \mid r) \cdot P(r)}{P(u)} = \frac{P_S(u \mid r) \cdot P(r)}{\sum_{i=1}^{n} P_S(u \mid r_i) \cdot P(r_i)}, \tag{B4}$$

where $P_S(u \mid r)$ is the likelihood with which the speaker produces utterance u when intending to communicate r, and $P(r)$ is the prior probability of choosing referent r. The denominator sums over the possible referents r_i.

Because a pragmatic listener considers what a speaker would say to be informative to a literal listener (and how costly that utterance would be), the pragmatic listener will make different inferences from an utterance than the literal listener. The literal listener can only draw on the prior $P(r)$ to distinguish between the two referents upon hearing a periphrastic causative as an utterance (since that utterance is true for both the direct and the non-direct cause). The pragmatic listener reasons that a speaker would choose the lexical causative if she wanted to refer to the direct cause, and so makes the inference that she probably wanted

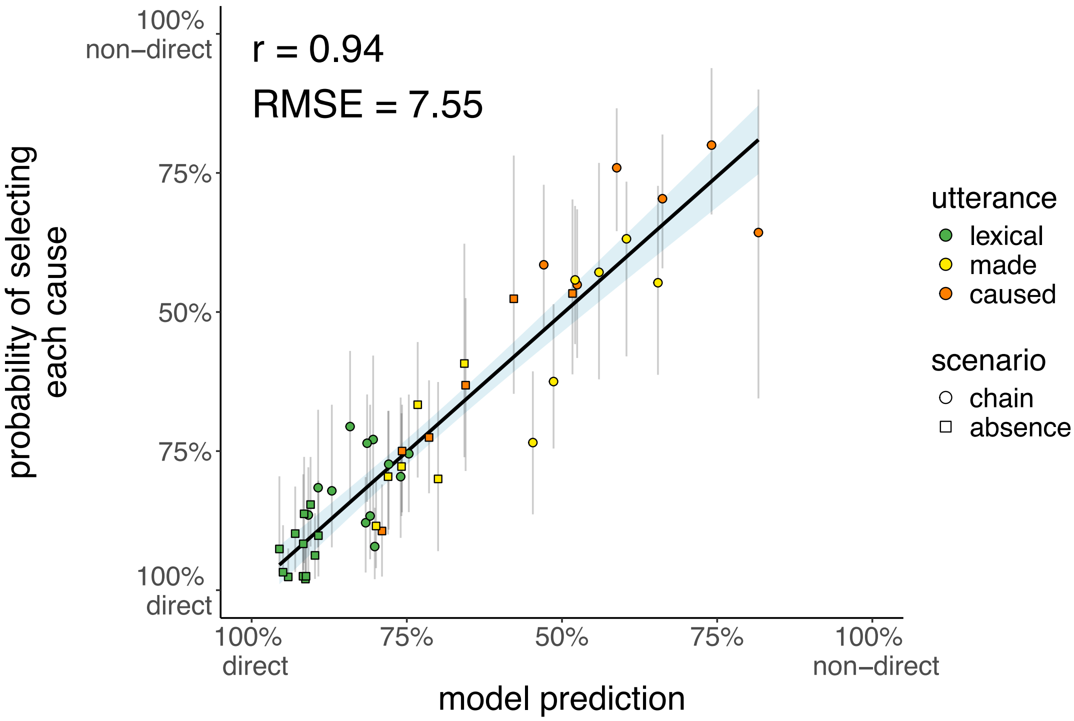

Figure B3

Relationship between model prediction and the probability with which participants selected the direct or non-direct referent. Colors indicate what utterance was used, and shapes indicate the scenario type. Each point indicates the proportion of children in a given age group who selected the direct or non-direct cause. Error bars show 95% bootstrapped confidence intervals. The blue ribbon shows the 95% confidence interval of the regression line.

to refer to the non-direct cause when she used a periphrastic causative. The strength of this inference increases the more costly the periphrastic causative is (relative to the lexical causative), and the more optimal the speaker is (i.e., the more likely they are to choose a more informative utterance).

Fitting the model to our data

The model assumes that the speaker optimality α increases with age (Figure 5A). This means that older children are more likely to think that a speaker would choose a more informative utterance. This assumption is consistent with other work which finds that the actions of older children are better explained by assuming that they behave more optimally and less noisily (Amemiya, Heyman, & Gerstenberg, 2024; Giron et al., 2023). The model assumes that the prior probability of choosing a non-direct referent is lower than choosing a direct referent, and that choosing to refer to an "absent cause" is particularly unlikely (Figure 5B). Finally, the model assumes that the cost of using a periphrastic expression increases with age, and that the cost of using "caused" is greater than the cost of using "made" (Figure 5C). Here, the cost for "lexical" is assumed to be 0 (as only the relative difference in cost matters). That the cost parameter increases with age means that children are assumed to develop a better appreciation that "caused" or "made" are less common utterances than lexical causatives, and that "caused" is more unusual than "made".

So, the model that captures developmental change in causal language understanding has 8 free parameters overall: two parameters that characterize how speaker optimality α changes with age (one parameter for the intercept and one for the slope; Figure 5A), two parameters that capture the prior $P(r)$ of referring to the non-direct referent in the chain and absence scenario (Figure 5B), and two parameters each (one for the intercept and one for the slope) for how the utterance cost c of "caused" and "made" changes across age (Figure 5C). We fitted the model by minimizing the sum of squared error between model predictions and averaged responses per age group (Figure 5A–C show the parameter fits). We had 12 data points per panel in (6 age groups times two utterances), so 48 data points in total. Figure B3 shows the model predictions together with the data. The model captures participants' responses well with $r = 0.94$ and RMSE $= 7.55$.

This model assumes that the meaning of the different causal expressions is fixed and doesn't change across development. While the results here show that it's possible to accurately model the developmental trends this way, we acknowledge that learning the meaning of the causal expressions and what their possible referents are is certainly part of a full developmental story (see Bergen, Levy, & Goodman, 2016; Bohn, Tessler, Merrick, & Frank, 2022, for work that models learning word meanings by pragmatically resolving lexical uncertainty).

References

Amemiya, J., Heyman, G. D., & Gerstenberg, T. (2024). Children use disagreement to infer what happened. *Cognition*.

Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, *9*, 20–1.

Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2022). Predicting pragmatic cue integration in adults' and children's inferences about novel word meanings. *Journal of Experimental Psychology: General*, *151*(11), 2927-2942.

Degen, J. (2023). The rational speech act framework. *Annual Review of Linguistics*, *9*(1), 519–540.

Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., & Goodman, N. D. (2020, jul). When redundancy is useful: A bayesian approach to "overinformative" referring expressions. *Psychological Review*, *127*(4), 591–621. Retrieved from `https://doi.org/10.1037%2Frev0000186` doi: 10.1037/rev0000186

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in exploration resemble stochastic optimization. *Nature Human Behaviour*, *7*(11), 1955–1967.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829. Retrieved from `https://doi.org/10.1016%2Fj.tics.2016.08.005` doi: 10.1016/j.tics.2016.08.005

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts.* New York: Wiley.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction.* Cambridge University Press.

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Computational Biology*, *4*(12), e1000254.