Developmental Psychology

Word Learning Challenges Explain Nonadult Possibility Language Comprehension in Preschoolers

Ailís Cournane, Anouk Dieuleveut, Chiara Repetti-Ludlow, and Valentine Hacquard

Online First Publication, September 1, 2025. https://dx.doi.org/10.1037/dev0002061

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

https://doi.org/10.1037/dev0002061

Word Learning Challenges Explain Nonadult Possibility Language Comprehension in Preschoolers

Ailís Cournane[1], Anouk Dieuleveut[2], Chiara Repetti-Ludlow[3], and Valentine Hacquard[4]

[1] Department of Linguistics, New York University
[2] Department of Linguistics, University of Geneva
[3] Neuroscience Institute, Carnegie Mellon University
[4] Department of Linguistics, University of Maryland, College Park

This article presents two experiments testing English children's understanding of the "force" of modals, asking whether they understand that *can* expresses possibility and *have_to* expresses necessity. Prior studies show that children tend to over-accept necessity modals in possibility situations and argue this behavior stems from conceptual difficulties reasoning about open possibilities. However, these studies typically test modal force using epistemic modality (knowledge-based), which is less input-frequent than nonepistemic modalities (actual-world priorities or goals) and involves speaker perspective-taking. Our results with more familiar teleological (goal-oriented) modality show that preschoolers have an adult-like understanding of possibility *can*, but they seem to treat necessity *have_to* as a possibility modal, in affirmative (Experiment 1) and arguably in negative sentences (Experiment 2). We take these systematic errors to call into question conceptual accounts. We argue that younger preschoolers' difficulties with modal force are due to word-learning challenges: They treat necessity modals as possibility modals.

Public Significance Statement

Children under age 5 seem to understand the word *can* much like adults do—but often treat *have to* as just another way to say *can*. Recognizing this language gap can support more developmentally appropriate communication with young children. Children under 5 might not fully grasp the sense of requirement in *have to*, which could affect how they respond to rules or directions.

Keywords: modality, possibility reasoning, possibility language, premature closure, word learning

In languages like English, modals are used to express various kinds of possibilities and necessities. In (1), the modals *might* and *must* express epistemic modality: possibilities and necessities given knowledge or evidence. The possibility claim in (1a) leaves other hiding places for Nick open besides the mentioned red box, while (1b) expresses the necessity that he is in the red box. In (2), *can* and *have_to* express teleological modality: possibilities and necessities given a particular goal. In (2a), going down the red road is one

Maria Arredondo served as action editor.

Ailís Cournane https://orcid.org/0000-0002-4288-4049

Additional online material on the Open Science Framework "Modal Word Learning Challenges" is available at https://doi.org/10.17605/OSF.IO/Q4G3N and https://osf.io/q4g3n/.

Valentine Hacquard and Ailís Cournane received funding from Grant BCS-1551628 from the National Science Foundation.

This work would not have been possible without all the families and participants, to whom the authors extend warmest thanks. The authors also thank *Carousel of Languages* in New York City and *Children Helping Science* (Massachusetts Institute of Technology) for invaluable help recruiting. Also, to our excellent research assistants for recruiting and testing: Sasha Frangulov ("Logan"), Madisen Fong ("Lily"), Mina Hirzel (Art), Yasmeen Carneiro, Madisen Fong, Hannah Mattis-Roesch, Sam Beames, Michael Marinaccio, Daniella Presti, Mark Bacon, Michelle Golden, Laura Phelan, Sofia Gonzalez, Vishal Sunil Arvindam, Sam Mitchell, Melissa Rojas, Adam Bell, Charlotte McFarland, and Liliana Que. For helpful insights and discussion, the authors thank Maxime Tulling, Sarah Phillips, and Brian Leahy. The authors also extend thanks to audiences at Generative Approaches to Language Acquisition, North American 9 (University of Iceland), University of Massachusetts at Amherst Linguistics, the University of Pennsylvania's Integrated Language Sciences and Technology, McGill Linguistics, the Massachusetts Institute of Technology Acquisition Group, the Child Language Lab at New York University and the Modality Group at the University of Maryland.

Ailís Cournane played a lead role in visualization and writing–original draft, a supporting role in software and validation, and an equal role in conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, and writing–review and editing. Anouk Dieuleveut played a supporting role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, visualization, and writing–original draft and an equal role in writing–review and editing. Chiara Repetti-Ludlow played a lead role in data curation and software, a supporting role in investigation, methodology, resources, supervision, validation, visualization, writing–original draft, and writing–review and editing, and an equal role in formal analysis and project administration. Valentine Hacquard played a lead role in validation and writing–review and editing, a supporting role in formal analysis, project administration, resources, supervision, and writing–original draft, and an equal role in conceptualization, funding acquisition, investigation, and methodology.

Correspondence concerning this article should be addressed to Ailís Cournane, Department of Linguistics, New York University, 10 Washington Pl, New York, NY 10003, United States. Email: cournane@nyu.edu

possible way—among others—for Cat to get to the store, while (2b) expresses the necessity of going down the red road (it is the only possibility).

(1) a. Nick might be hiding in the red box. EPISTEMIC POSSIBILITY
b. Nick must be hiding in the red box. EPISTEMIC NECESSITY
(2) a. Cat can go down the red road to get to the store. TELEOLOGICAL POSSIBILITY
b. Cat has to go down the red road to get to the store. TELEOLOGICAL NECESSITY

Previous studies primarily use epistemic modality to test children's comprehension of modal "forces" (possibility vs. necessity). Results show that children display nonadult behaviors on modal comprehension tasks until age 6. Children tend to over-accept possibility modals (e.g., *might, can*) in situations of necessity (Noveck, 2001; Ozturk & Papafragou, 2015). For instance, in a context where Nick is hiding in a red or a blue box, but the blue box is revealed to be empty, children readily accept possibility sentences like (1a) to describe the situation, while adults prefer (1b). Children also tend to over-accept necessity modals (*must, have_to*) in contexts where several possibilities remain open. For instance, in a context where the evidence leaves open whether Nick is in the red or blue box, adults only accept possibility modals (1a), but children also accept necessity modals as in (1b; Bleotu et al., 2021; Moscati et al., 2017; Ozturk & Papafragou, 2015). These nonadult behaviors have been argued to arise from either (a) pragmatic immaturity, affecting their ability to judge the appropriateness of modal sentences, or (b) conceptual immaturity, affecting their ability to reason under indeterminacy. The most common conceptual immaturity view holds that children perform "Premature Closure" (Acredolo & Horobin, 1987), randomly discarding one open possibility in possibility scenarios, which leads them to accept necessity modals about half the time (Bleotu et al., 2021; Moscati et al., 2017; Ozturk & Papafragou, 2015). B. Leahy and Zalneiriunas (2021), by contrast, argue that children's nonadult-like modal behavior reflects a deeper conceptual deficit—namely, an inability to represent modal possibilities (B. P. Leahy & Carey, 2020), rather than difficulty maintaining them.

The fact that prior studies primarily rely on epistemic flavor to test children's force comprehension may have introduced additional conceptual difficulties over nonepistemic modalities, which could mask children's competence with basic modal concepts and vocabulary. Epistemic reasoning involves representing alternative belief or knowledge states (Phillips & Kratzer, 2024; Robinson et al., 2006). In prior epistemic tasks, children typically need to consider the speaker's perspective and knowledge state, a requirement known to complicate tasks for preschoolers (see Ünal & Papafragou, 2016). Furthermore, prior studies have tacitly assumed that children know the lexical meanings of the modals tested, both in terms of "flavor" (e.g., epistemic vs. teleological) and "force" (possibility vs. necessity). From this assumption, findings are thought to reflect possibility reasoning difficulties. However, children's nonadult behaviors could arise from a nonadult understanding of flavor or force.

First, children may not know the intended modal flavor, as the same modal words can express different kinds of possibility or necessity with differing truth conditions. If children understand the modals tested in ostensibly epistemic scenarios as expressing deontic modality, they could end up accepting a sentence like (1b) because, for all they know, Nick may be *obligated* to be in the red box (even if epistemically there are other possibilities open), a real possibility, given that nonepistemic modal uses typically occur at least 10 times more frequently than epistemic uses in children's input cross-linguistically (Cournane & Tailleur, 2020; van Dooren et al., 2019, 2022; Veselinović & Cournane, 2020).

Second, children may not know the force of modals, particularly that *must* and *have_to* express necessity rather than possibility. Modals pose a classic "subset problem" (Dieuleveut et al., 2022; Gualmini & Schwarz, 2009; Piantadosi, 2011; Xu & Tenenbaum, 2007, i.a.) because necessity entails possibility: situations of necessity are also situations of possibility. This could lead children to lexicalize the possibility meaning for necessity modals. If so, we would expect them to accept those modals in possibility situations. This nonadult force analysis for necessity modals would have further consequences for their possibility modals: If children lack necessity modals, then they should accept possibility modals in necessity situations for lack of a more suited, stronger alternative (put another way, they cannot make scalar implicatures if they do not know a stronger dual). In fact, this is what speakers of a language like Nez Perce do, as their language lacks a necessity modal (Deal, 2011).

In this article, we explore whether preschoolers know the force (possibility vs. necessity) of modals *can* and *have_to* in teleological scenarios. We side-step possible additional difficulties stemming from epistemic modality by using the more familiar and, arguably, conceptually simpler, teleological modality (1a,b). By using this very imageable flavor of nonepistemic modality (cf. deontic modality, Cummins, 1996; Dack & Astington, 2011; Hirst & Weil, 1982), and the two modal verbs English-learning children have the most input experience with (van Dooren et al., 2022), we give children a better chance to show what they know about force than prior studies. We test preschoolers (aged 3 to 5) and adult controls in a series of truth value judgement (Crain & Thornton, 1998; Experiment 1) and sentence choice (Experiment 2) tasks with *can* and *have_to*, and negated *can't* and *doesn't have_to*, in teleological possibility, necessity and impossibility scenarios. In contrast to prior studies testing negated modals (Bleotu et al., 2023; Gualmini & Moscati, 2009; Koring et al., 2018, Moscati & Crain, 2014, Moscati et al., 2016), we use negation to assess whether children's response patterns align with our hypothesis for nonadult modal force, especially for *have_to*. By reversing the truth conditions in our scenarios on assumptions of adult-like scope for possibility modals, negation allows us to explore child modal force more precisely, offering more insights into children's basic lexical knowledge of modals.

Our results are consistent with our hypothesis that children do not have adult-like modal word knowledge. From age 3 children know that *can* is a possibility modal, as their comprehension behaviors are adult-like in all scenarios, and they readily differentiate positive and negated sentences in an adult-like way. However, children do not treat *have_to* as a necessity modal until past age 5, when some are beginning to behave adult-like. Three- and 4-year-olds' comprehension behaviors are consistent with *have_to* expressing possibility. Critically, in contrast to prior tasks, responses to sentences with necessity modals in possibility scenarios are not at chance but systematically pattern with possibility modal *can* (see also B. Leahy & Zalneiriunas, 2021). Our results therefore argue against the idea that

preschool children toss out one open possibility at random in possibility scenarios (Premature Closure), at least in root modalities. Results are consistent with our hypothesis that longstanding preschooler difficulties with modal force comprehension result from systematic lexical issues: while preschool-aged children have worked out the force of possibility modals like *can*, they lack an adult-like understanding of the force of necessity modals like *have_to* through at least age 5.

Modal Acquisition: Learning Challenges and Existing Results

In the developmental literature, there is an important debate about whether children are born with the ability to represent distinct open possibilities or whether this ability develops over time. From infancy, children seem able to consider different possible futures, and anticipate probabilistic outcomes (Alderete & Xu, 2023; Cesana-Arlotti et al., 2018, 2022; Téglás & Bonatti, 2016; Téglás et al., 2007), suggesting that these modal concepts are in place early. But children also seem to have difficulty acting upon possible futures until age 4. In Redshaw and Suddendorf's (2016) forked-tube task, participants had to catch a ball dropped from a tube with two openings at the bottom: while adults and 4-year-olds successfully placed a hand under each opening to guarantee catching the ball, 3-year-olds only covered one opening at a time. Redshaw and Suddendorf attributed this behavior to an inability to simultaneously represent alternative outcomes (see also B. P. Leahy & Carey, 2020; Mody & Carey, 2016). Proponents of the view that modal representations may be in place early blame this behavior on additional task demands (e.g., Turan-Küçük & Kibbe, 2024). But, all accounts agree that children can represent multiple possible future outcomes by age 4, though B. Leahy, and Zalneiriunas (2021) suggested that some children may continue to struggle with representing modal possibilities beyond this age.

Despite basic conceptual underpinnings likely being in place by age 4, studies looking at modal force comprehension (e.g., *may* vs. *must*) report nonadult behaviors up to at least age 6. In their groundbreaking study, Hirst and Weil (1982) tested modal contrasts (e.g., *may* vs. *must*) in both epistemic (hidden object search task) and deontic (rules from teacher puppets) act-out scenarios, for example, prompting children to find a peanut based on two statements: *The peanut may be under the box* versus *the peanut must be under the cup*. Their deontic results were surprisingly poor and lagging behind epistemic results, especially considering production data which shows more deontic use earlier than epistemic (Cournane, 2021; Hickmann & Bassano, 2016; Papafragou, 1998). This difference was attributed to the convoluted nature of their deontic task and contributed to the subsequent focus on epistemic modality to test modal force. Their epistemic results, and those of subsequent studies, show that by 5-years-old children typically show adult-like preferences for acting on the stronger statement (e.g., searching under the cup for the peanut, as *must* is stronger than *may;* Bascelli & Barbieri, 2002; Byrnes & Duff, 1989; Hirst & Weil, 1982; Ozturk & Papafragou, 2015, Experiment 2), but still struggle with necessity, especially compared to plain assertions (*must* vs. *is*).

In truth-value judgement tasks where children judge one modal sentence at a time, preschool children show two consistent patterns across all tasks; they both over-accept possibility modals in necessity scenarios (Noveck, 2001; Ozturk & Papafragou, 2015) and necessity modals in possibility scenarios (Moscati et al., 2017; Ozturk & Papafragou, 2015). Children's over-acceptance of possibility modals in necessity contexts may be a pragmatic issue: A weaker statement like (1a) is logically true in a necessity context, but it is judged inappropriate when the more informative stronger statement (1b) is also true and relevant. When a necessity statement would be relevant, the use of a weaker statement typically triggers a scalar implicature that the stronger statement does not hold, as the speaker would otherwise have used it (Grice, 1975): "Nick *might* be in the red box" implies that "Nick *doesn't have_to* be in the red box." Noveck (2001) suggested children make judgements based only on the logical meaning of modals. Various studies since show that children's difficulty with implicatures is not with the inference process per se but with knowing when a stronger alternative is relevant (see Barner & Bachrach, 2010; Barner et al., 2011; Chierchia et al., 2001; Skordos & Papafragou, 2016). Results from scalar implicatures more generally show that children perform better when alternatives are provided (e.g., Chierchia et al., 2004; cf. Hirst & Weil, 1982). Thus, children's over-acceptance of possibility modals in necessity scenarios may be due to not knowing when a necessity modal would be more appropriate and relying on logical meanings alone. Ozturk and Papafragou (2015) showed that adults also readily accept possibility modals in necessity contexts, unless an alternative necessity statement is made salient, suggesting that children may not differ from adults with *can* or *may*.

Children's over-acceptance of necessity modals is more puzzling because necessity modals are logically false in possibility situations: if there are two possible spots where Nick could be hiding, then (2b) is false. In a representative study, Ozturk and Papafragou's (2015) Experiment 1 showed children animated stories where an animal would hide in a box. The experimenter would then either make a modal statement (e.g., "The X has to be in the blue box") or ask a question (e.g., "Does the X have to be in the blue box?"), in four different story types: Story 1 (Necessity; single closed box); Story 2 (Necessity; 2 boxes: mentioned box closed, other box empty); Story 3 (Possibility: two closed boxes); and Story 4 (Impossibility; 2 boxes: mentioned box empty). Children and adults tended to accept possibility modals in both possibility and necessity contexts. Adults tended to accept necessity modals in necessity contexts and reject them in both possibility and impossibility contexts. Children also accepted necessity statements in both necessity contexts (though they were at chance in the question condition in Story 2) and rejected them in impossibility contexts. Crucially, however, children accepted necessity modals in possibility situations half the time. Note that because results were reported in the aggregate, this 50% behavior may have come from a single population or a bimodal distribution.

Moscati et al. (2017) used a 3-box paradigm (adapted from Noveck, 2001), to collect combined eye-tracking and judgement experiment with younger 5-year-olds. Subjects saw the contents of two boxes and heard statements about what is in a third closed box which matches one of the open boxes. For example, if both open boxes contain a parrot, then the closed box *must* also, but if only one open box contains a parrot, then the closed box only *might* have one. Children had difficulty making judgements in undetermined scenarios (with open possibilities) over determined ones (no

uncertainty, akin to necessity). Children did poorly for both *might* and *must* sentences. However, implicit eye-tracking measures showed children looked more at the unmentioned item from the visual scene *must* sentences than *might*, like adults (although more briefly). This suggests that 5-year-olds differentiate the modal meanings in the right direction but have trouble with the truth value judgment, which Moscati et al. (2017) attributed to nonlinguistic difficulty with uncertainty judgements.

Ozturk and Papafragou (2015) and Moscati et al. (2017) attributed the over-acceptance of necessity modals in possibility scenarios to conceptual difficulty with reasoning about open possibilities. According to what Acredolo and Horobin (1987) coined "Premature Closure," young children have trouble reasoning over indeterminate states of affairs, and resolve this difficulty by ignoring some of the open possibilities. While in toddlers the ability to mentally represent more than one open possibility may arguably be absent (see B. P. Leahy & Carey, 2020), preschoolers may be only able to briefly maintain such representations. In the above tasks they arbitrarily eliminate some alternatives (e.g., one of the hiding spots). When children randomly toss out one possibility, a necessity statement ends up being true half the time (when the child tosses out the possibility of the unmentioned box). This propensity also occurs in nonlinguistic tasks (Beck & Robinson, 2001; Robinson & Whittaker, 1986; Robinson et al., 2006; Somerville et al., 1979; Taylor, 1988).

Whether such a propensity can really explain children's difficulty with epistemic modal judgments is not entirely clear. If premature closure is at play, the same difficulty should occur for possibility modals as for necessity modals (Moscati et al., 2017). We also would expect children to only accept possibility statements in possibility contexts half the time when children toss out the unmentioned box, but existing results are mixed. While Moscati et al. (2017) found that children were nonadult in possibility scenarios for both *might* and *must*, Ozturk and Papafragou (2015) found that children accepted *may*-statements at adult-like rates but responded "yes" 44% of the time to *can*-questions. Ozturk and Papafragou avoided *may* for the question condition due to unnatural phrasing in American English, but *can* is not used epistemically except when negated (Iatridou & Zeijlstra, 2013), making the question incompatible with its intended meaning. Adult participants may have accommodated, by interpreting the question as "is it circumstantially possible?" while children, possibly viewing the experimenter as an authority, may have been inclined to interpret it as "is it allowed?" This is perhaps especially likely since the experimenter was an adult who also gave the rules of the task. This could explain these variable responses. Moscati et al.'s (2017) chance-level responses with *might* in undetermined scenarios provide strong evidence for under-acceptance of possibility modals in epistemic contexts, supporting their conclusion that premature closure is at play. Bleotu et al. (2021) reported similar findings with Romanian epistemic adverbs. Since both studies involved 5-year-olds, inability to represent modal possibility seems unlikely—this is typically proposed for children under 4 (B. P. Leahy & Carey, 2020)—and is further contradicted by Moscati et al.'s eye-tracking data showing that 5-year-olds distinguish *might* from *must*.

Children's difficulties with open possibilities may last longer with epistemic reasoning. Robinson et al. (2006) tested 4- to 6-year-olds in a task where they had to place trays to catch colored blocks. The experimenter would draw blocks from one of two bags, one containing black blocks, and the other containing a mix of orange and green blocks. The experimenter would then place the block behind a cardboard wall and push it through a door whose color matched that of the block. As in the forked tube task, the question was whether children would anticipate two distinct outcomes when the experimenter picked a block from the orange and green bag and place trays under both doors. Trials critically differed in when children had to place the trays. In the "physical uncertainty" condition (corresponding to nonepistemic modality), children had to place trays before the block was picked. In the epistemic condition, the experimenter would place the block behind the door (unseen to the child), and only then ask the child to place the trays: In this condition, the type of block was physically determined, but still epistemically uncertain for the child. The 4- to 5-year-olds succeeded in placing two trays in the physical uncertainty condition but failed to do so in the epistemic condition. The 5- to 6-year-olds performed better but still struggled. This asymmetry between nonepistemic and epistemic reasoning may be due to additional difficulty with the metarepresentational aspects involving epistemic states (Phillips & Kratzer, 2024).

B. Leahy and Zalneiriunas (2021) argued that children's non-adult-like modal behavior reflects a conceptual deficit, building on B. P. Leahy and Carey's (2020) proposal that young children cannot represent modal possibilities and instead rely on minimal representations of possibility—simulating a single outcome rather than multiple possibilities. They tested children aged 4 to 7;11 and adults on *can* and *have_to* questions using single and forked tube tasks (adapted from Redshaw & Suddendorf, 2016). These scenarios target future-oriented possibility and necessity, making *can* and *have_to* well-suited for testing (we consider these root circumstantials). Both a linguistic task (If I drop a marble in here, can it come out here?) and a nonlinguistic task (placing a wagon to catch the marble) were used (note the latter involves potential action-planning confounds, see Turan-Küçük & Kibbe, 2024). *Can* responses were adult-like from age 4, but 4- and 5-year-olds treated *have_to* similarly to *can*, giving the same response patterns. Leahy and Zalneiriunas argue that many 4- and 5-year-olds still lack modal representations of possibility. By-individual assessments and the nonlinguistic task suggest that *some* have acquired this ability, while others have not. They argue that young children may show "accidental" success with *can*, as checking one path can mimic an adult response. *Have_to* responses may similarly reflect single-shot reasoning, leading them to pattern with *can*. Relying on minimal representations of possibility, children fail to distinguish possibility from necessity—and thus *can* from *have_to*.

We suggest children's difficulties with modal force tasks may alternatively stem from lexical challenges, given the complex, language-specific, form-meaning mappings of modal words (see also Cournane, Hirzel, & Hacquard, 2023). In languages like English, the same modal words can express different "flavors." For example, *have_to* can express an epistemic necessity or various types of nonepistemic necessities, like deontic or teleological. English-learning children start producing modal verbs early, first with nonepistemic flavors around age 2, followed by epistemic flavors about a year later.[1] Learning that modals like *have_to* can express epistemic flavors may be particularly challenging, given that nonepistemic flavors are much more frequent than epistemic in child

[1] Cournane (2021) noted that epistemic adverbs like *maybe* are produced from age 2, so this flavor asymmetry is not readily explained by a conceptual gap, as commonly argued (Papafragou, 1998; Sweetser, 1990).

directed speech (van Dooren et al., 2022). We suggest the possibility that in prior epistemic tasks, children accessed nonepistemic interpretations of the modals tested, as those meanings may be more familiar to young children, that is, (1b) could be interpreted as deontic. Flavors differ in truth conditions and predicted responses, as it may be epistemically *possible* that Nick is in one of two boxes, but deontically *necessary* (given some perceived authority, likely the experimenter) that he be in the mentioned box.

Alternatively, children's difficulties with force tasks may arise from not yet knowing the lexical force of the modal words tested. English distinguishes modal force lexically, with both possibility (*can, may*), and necessity modals (*have_to*, *must*). Figuring out the force that each modal expresses comes with its own challenges, as modals present a "subset problem." The sentence in (3) entails the one in (4), when keeping flavor constant. This means that, whenever (3) is true, (4) is also true, making it possible for learners to wrongly assume that *have_to* merely expresses possibility.

(3) Cat has to eat fish.

(4) Cat can eat fish.

Some prior studies have looked at modal force in natural and elicited production. These studies provide initial support for the hypothesis that young children struggle with the force of necessity modals in particular. Dieuleveut et al. (2022) compared modals naturally produced by 2-year-olds and their mothers. To assess how adult-like children's modal productions were, Dieuleveut et al. ran a *Human Stimulation Paradigm* experiment (Gillette et al., 1999), in which adult participants had to guess the force of modals produced by children and their mothers, by reading snippets of real conversations in which the modals had been blanked out. Participants were well above chance at guessing the correct force of the mothers' possibility and necessity modals and that of children's possibility modals, but not children's necessity modals. Thus, by age 2, children seem to be using possibility modals in an adult-like way, but not necessity modals. In an elicited production task, Cournane, Hirzel, and Hacquard (2023) had 3- and 4-year-olds produce modals of their own choosing to describe epistemic and teleological possibilities and necessities, by providing a scenario and sentence frame with the modal obscured (e.g., "Nick <*pink noise*> be hiding in the red box"). Results show that children prefer using possibility modals across both force situations and both flavors and that when they do use necessity modals, they use them indiscriminately in both necessity and possibility situations.

Table 1 summarizes prior results for modal force for positive modal sentences. In sum, they show that children are nonadult in their use and comprehension of modals in the preschool years, but the source of these difficulties is unclear. Do children's nonadult behaviors reflect difficulty reasoning about open possibilities, or are they linguistic in nature? Do children know the flavor and force of modal words? Are their difficulties specific to epistemic modality and reasoning, or are they more general? To add to our understanding of both modal force comprehension and possibility reasoning development, our experiments test the comprehension of modals *can* and *have_to* in teleological modality, with preschoolers and adult controls. We use *can* and *have_to* because they are the most frequently used possibility and necessity modals, respectively, in the input to young English-learning children, and both are frequently used in teleological modality in the input (van Dooren et al., 2022).

Another way to assess the force children assign to modals is by examining how they interpret modals in the presence of negation. In production, negation appears reliably by age 2 across various morphological forms in English (e.g., *no*, *not*, *n't*), and its usage is largely consistent with adult-like patterns (Jasbi et al., 2021; see also Bloom, 1970; Choi, 1988; Pea, 1980). Denial uses of negation—most relevant for propositional reasoning and the focus of the present study—emerge consistently in form by age 2 in many children and reach adult-like rates by age 3 (Jasbi et al., 2021). In comprehension, children can use negation productively in word learning at age 2 (de Carvalho et al., 2021, 2023) and reliably understand it by age 3 (Reuter et al., 2018).

Table 1
Summary of Stimuli and Results for Prior Modal Force Comprehension Studies

Study	$Age_{(M)}$	Flavor	Method	Modal(s)	Result
Noveck (2001), Experiment 1, youngest group	5;5	Epistemic	3 box paradigm TVJT	*might* *have_to*	Chance Chance
Gualmini and Moscati (2009)	4;5	Root (ability)	Story TVJT	*potere* (Italian "can")	Adult-like direction
Moscati and Crain (2014)	5;4	Epistemic[a]	3 box paradigm TVJT	*potere* (Italian "can")	Chance
Ozturk and Papafragou's (2015), Experiment 1	5;1	Epistemic[a]	2 box paradigm TVJT/Questions	may_{TVJT}/can_{Q} *have_to*	Adult-like/chance Chance
Moscati et al. (2017)	5;0	Epistemic	3 box paradigm Eye tracking/TVJT	*might* *must*	Adult-like (eye) Chance (TVJT)
Koring et al.'s (2018), younger group	5;5	Epistemic[a]	3 box paradigm TVJT	*moeten* (Dutch "must")	Chance
Bleotu et al. (2021)	5;6	Epistemic	Shadow paradigm	*poate* (Romanian "maybe") *sigur* (Romanian "certainly")	Adult-like Chance
B. Leahy and Zalneiriunas (2021)	4;6 5;2	Root (circumstantial)	One and two tubes task Questions	*can* *have_to*	Adult-like Nonadult-like

Note. Positive scenarios only. TVJT = Truth Value Judgement Task.
[a] Denotes that deontic interpretations were also possible for the modals tested.

Several studies examine children's comprehension of how modals scope relative to negation in English, Italian, Dutch, and Romanian (Bleotu et al., 2023; Gualmini & Moscati, 2009; Koring et al., 2018, Moscati & Crain, 2014, Moscati et al., 2016). This is a complex topic, since possibility and necessity modals vary with respect to whether they scope over or under negation depending on force and flavor interpretation, Iatridou and Zeijlstra (2013), see Dieuleveut (2021) for discussion of issues pertaining to their acquisition. For instance, in English, necessity modals like *must* take a wide scope with respect to negation (*must not* = necessary > not), but modals like *have_to* and *need_to* take a narrow scope (*doesn't have_to* and *need not* = not > necessary).

Overall, these studies show that children tend to interpret possibility modals as scoping under negation ("not possible"), but necessity modals as scoping over negation ("necessary not"), regardless of what they do in the adult grammar across different languages. For instance, children tend to reject "There might not be a horse in the box" (possible > not) when it is possible that there is no horse in the box (see Bleotu et al., 2023; Gualmini & Moscati, 2009; Koring et al., 2018; Moscati & Crain, 2014; Moscati & Gualmini, 2007; Noveck, 2001). These studies generally take for granted that children know the underlying force of their modals, to focus on what they know about their scope relative to negation (i.e., whether children choose strong inverse scope interpretations rather than weak surface scope interpretations). One proposal for this preference for strong interpretations is the Semantic Subset Principle, a variant of the Subset Principle, according to which children acquire scope ambiguities by first assuming stronger meanings, regardless of their availability in the adult grammar (Crain, 2012; Crain & Thornton, 1998; Crain et al., 1994; Moscati et al., 2016; for critiques, see Musolino, 2006; Musolino et al., 2019).

Note, however, that outside of modals, children tend to assume "surface scope," to a fault (Lidz & Musolino, 2002). An alternative interpretation of these results consistent with this tendency would be that children assume that modals uniformly scope under negation, but they tend to interpret necessity modals as possibility modals (not possible = necessary not).

In this article, we explicitly test for children's force comprehension and further look at how they interpret these modals with and without negation. In the adult grammar, the modals we use (*can* and *have_to*) are well-behaved in their scopal relations to negation: both scope under negation. If children treat *have_to* sentences as expressing possibility in positive contexts, but impossibility in negative contexts, this will argue that they have the correct scope interpretation but the wrong modal force interpretation, contrary to the interpretation of the previous results on children's interpretation of modals with negation.

From this background, we focus on the following hypotheses.[2]

1. PREMATURE CLOSURE HYPOTHESIS: Children have difficulty reasoning about open possibilities, but know the force of the modals tested. They reduce uncertainty by randomly eliminating one open possibility before judging the modal statement. This predicts children to over-accept necessity modal *have_to* in teleological possibility situations about half the time, as in previous epistemic studies. It also predicts children to over-reject possibility modal *can* half the time, as children face the same uncertainty in possibility situations.
2. NONADULT LEXICAL FORCE HYPOTHESIS: By age 4, children can conceptually handle open possibilities, but lexical issues persist. Given the directionality of the subset problem for modals, we expect children to struggle particularly with the force of necessity modals. If children treat *have_to* as a possibility, they should accept it in possibility scenarios.

Experiment 1 is a truth value judgement task, inspired by Ozturk and Papafragou (2015), in which participants judge statements in three scenarios (POSSIBILITY, NECESSITY, and IMPOSSIBILITY). Experiment 2 tests modals with negation (*can* vs. *can't*; *has_to* vs. *doesn't have_to*), using a sentence choice task.

Transparency and Openness

All data, materials, and analysis code are available on the Open Science Framework at https://osf.io/q4g3n/. This study was not preregistered.

Experiment 1

Method

Participants

Participants were 24 monolingual English preschool children, recruited and tested through daycares in the New York City greater metropolitan area. An additional four children were excluded from the study because they were non-English-dominant bilinguals. The 24 children ($n_{female} = 9$, $n_{male} = 15$) included in the analysis ranged in age from 3;0 to 5;2 (M_{age} = 4;1, SD_{age} = 0.57). We did not conduct an a priori power analysis, but rather based our sample sizes on prior similar studies, with consideration of the number of participants per between-subjects condition and the number of test observations per participant (Moscati et al., 2017; Noveck, 2001; Ozturk & Papafragou, 2015). Our sample is younger than in previous epistemic studies because children at this stage have been reliably producing both possibility and necessity modals for over a year in root flavors (*can* and *have_to* being the most common; Cournane, 2021; van Dooren et al., 2022) and are beginning to pass nonlinguistic tasks about open possibilities (summarized in B. P. Leahy & Carey, 2020). Adult controls were native speakers of English, tested in New York City ($n = 24$, $n_{female} = 12$, $n_{male} = 12$; M_{age} = 23 years, $SD_{age} = 3.78$, $Range_{age}$ = 19–35). Parental consent and child assent were gathered prior to testing children, and adults consented prior to testing.

We sampled participants based on language eligibility criteria, monolingual American English, defined as no or minimal (<20%) use or direct input from other languages. The sample varies across different racial and ethnic groups, but specific ethnicity or race demographics were not collected. The specific child samples by-modal were *can*: ($n = 12$, $n_{female} = 6$, $n_{male} = 6$; M_{age} = 4;0, SD_{age} = 6.45, $Range_{age}$ = 3;0–4;11), *have* ($n = 12$, $n_{female} = 3$, $n_{male} = 9$; M_{age} = 4;2, SD_{age} = 7.20, $Range_{age}$ = 3;4–5;2).

[2] B. P. Leahy and Carey (2020) and B. Leahy and Zalneiriunas (2021) came out after we conceived of our study, therefore we did not directly consider their conceptual deficit proposal. We will instead consider it post hoc in the Discussion.

Figure 1

Introduction Slides (A, B) and Sample Training Item (C) From Experiment 1

Note. (A) Narrator: "We're going to listen to some stories about Cat and her friend Mouse." (B) Narrator: "Logan here [Logan: 'Hi, I'm Logan!'] is also going to listen to the stories and then tell us something about them. But he doesn't always pay attention. We need your help! Listen carefully to our stories and use the pictures to help Logan. Ready?". (C) Narrator: "Today, Cat will read a book. Cat wears glasses to help her see." Logan: "Cat is wearing glasses." Narrator: "Is Logan right?". See the online article for the color version of this figure.

Procedure and Experimental Design

At the beginning of the experiment, the narrator introduces the participant to Logan, the puppet who will say the test sentences, and to Cat, the main character of the story (Figure 1). The narrator establishes that Logan does not always pay attention ("silly puppet" paradigm, Crain & Thornton, 1998) and encourages the participant to help.

Four simple training items help garner both "yes (he's right)" and "no (he's wrong)" judgements from participants (Figure 1C). After the training items, a screen that says "Birthday Party!" appears and the narrator briefly introduces the overarching storyline of the main items (5).

(5) Narrator: "Cat is organizing a birthday party for her friend Mouse! She's going to buy some things for the party at different stores around town."

In the overarching story, Cat is organizing a birthday party for her friend, Mouse. Cat visits many shops and houses around town (e.g., the bakery, Dog's house). Each visit substory is a test item. The participant watches the story, with Logan onscreen, also listening along. Experimental lists were created in Keynote (Apple) and played as semianimated stories. Images were created in Draw (Adobe), and audio was recorded and cleaned using Praat (Boersma & Weenink, 2013). Three native speakers of American English provided voices for the narrator, Logan, and Lily. An experimenter was always present to moderate the experimental sessions. To mitigate against interpretation of our test sentences as deontic, we have child-coded puppets deliver the test sentences, instead of the experimenter (as in Ozturk & Papafragou, 2015, Experiment 1), to reduce the role of authority and rules (see Dack & Astington, 2011). We trained our voice actors to use a nonauthoritative sounding, somewhat dopey, prosody (in the spirit of the "silly puppet" paradigm). To support the teleological flavor, we stress the goal on every test sentence ("To get to the pizzeria,") and show no authority or knowledge about the answers Logan gives.

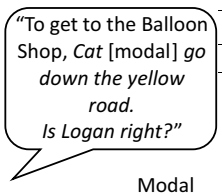
Test items involved two possible roads to get to a destination, highlighting the goal to support teleological modality interpretations. We manipulated two factors: *SCENARIO*: POSSIBILITY, NECESSITY, IMPOSSIBILITY, and *MODAL*: CAN, HAVE_TO (Table 2). For POSSIBILITY scenarios, both roads are open and allow Cat to reach the goal. For NECESSITY and for IMPOSSIBILITY scenarios, one of the roads is blocked with a construction barrier and cones. In NECESSITY scenarios, the audio stimuli mention the open road, in IMPOSSIBILITY scenarios, the audio mentions the blocked road. These stimuli are as simple or simpler than prior epistemic stimuli, as there is no need to infer from indirect evidence as with epistemic modality (Von Fintel & Gillies, 2010).

(6) Logan: "To get to the {store/house/ … }, Cat {can/has to} go down the [mentioned road]."

Table 2 provides the logical response for each modal by scenario. "Logical response" refers to what is predicted by the logic of

Table 2

Illustration of Scenarios (POSSIBILITY, NECESSITY, IMPOSSIBILITY) and Modals (CAN, HAVE), With Logical Responses to Prompt "Is Logan Right?" by Scenario and Modal



		Scenario		
	Modal	POSSIBILITY	NECESSITY	IMPOSSIBILITY
Logical Response	CAN	Yes	Yes	No
	HAVE_TO	No	Yes	No

Note. See the online article for the color version of this table.

possibility or necessity for the scenario. The crucial scenarios that distinguish between *can* and *have_to* sentences are POSSIBILITY scenarios: *can* sentences are true, *have_to* sentences are false. In NECESSITY scenarios, both *can* and *have_to* sentences are true, since necessity entails possibility. In IMPOSSIBILITY scenarios, both sentences are false since it is neither possible nor necessary to take the blocked road. While CAN is logically true in necessity scenarios and we predict a "yes" response, CAN is pragmatically underinformative (stronger HAVE_TO is better). However, we expect adults and children to respond based on the logical predictions, as in Ozturk and Papafragou (2015), Experiment 1, where in the parallel epistemic case adults accepted CAN at the ceiling (not generating scalar implicatures). Dieuleveut et al. (2021) further showed adults very rarely generate scalar implicatures for *can* in teleological scenarios like ours.

Figure 2 illustrates a sample item with *have_to* tested in a NECESSITY scenario. For each substory, the narrator first introduces the goal destination (e.g., the Balloon Shop; Figure 2A), each colored road (Figure 2B), and the construction blockage when present (Figure 2C). Logan then says the test sentence (6), and the narrator asks the participant if Logan is right (Figure 2D). Logan faces forward and his mouth moves when he speaks, but he looks at the inset image while the Narrator speaks.

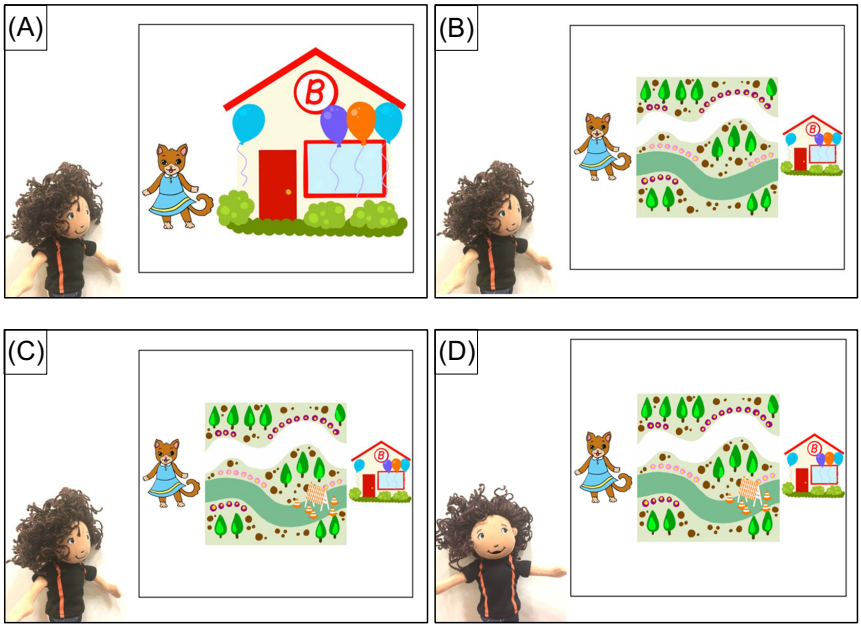
We varied *MODAL* (CAN, HAVE_TO) between subjects (12 participants per modal, per age group), and *SCENARIO* (POSSIBILITY, NECESSITY, IMPOSSIBILITY) within subjects. There were four CAN lists and four HAVE lists, with semirandomized orders of items for SCENARIO for each list (we generated random orders, but ensured no more than two items from the same condition appeared in a row). Three children and three adults were tested on each list. Every participant first saw four training items, to familiarize them with the task and make sure they could answer both "yes" and "no." To pass, children had to get three of four correct and answer both a correct "yes" and "no" item. All children included in our sample passed the training (nearly all with 4/4). Participants then saw 15 test items with five items from each scenario type (POSSIBILITY, NECESSITY, IMPOSSIBILITY). Order of mention of roads (top, bottom), mentioned color (red, green, etc.), and target road position (top, bottom) was fully counterbalanced. The dependent variable was yes/no judgement of whether Logan was right or not. We prompted about one third of the items (marked on scoresheets with an asterisk) with "How come?" for qualitative responses after participants made their judgements. The experimenter recorded all participant responses by hand on a scoresheet. Most sessions were recorded (when parental consent was obtained), allowing double-checking of responses as needed.

This study was not preregistered. All experimental item lists, data, code, full regression models, individual participant responses, and additional sample demographics (parental education levels, which overall skew heavily to college-educated) are available in our additional online material folder at https://osf.io/q4g3n/.

Results

Figure 3 plots "yes" responses to the prompt "Is Logan right?" according to which modal Logan used in the test utterance (*can*, *have_to*) and which scenario was depicted (POSSIBILITY, NECESSITY,

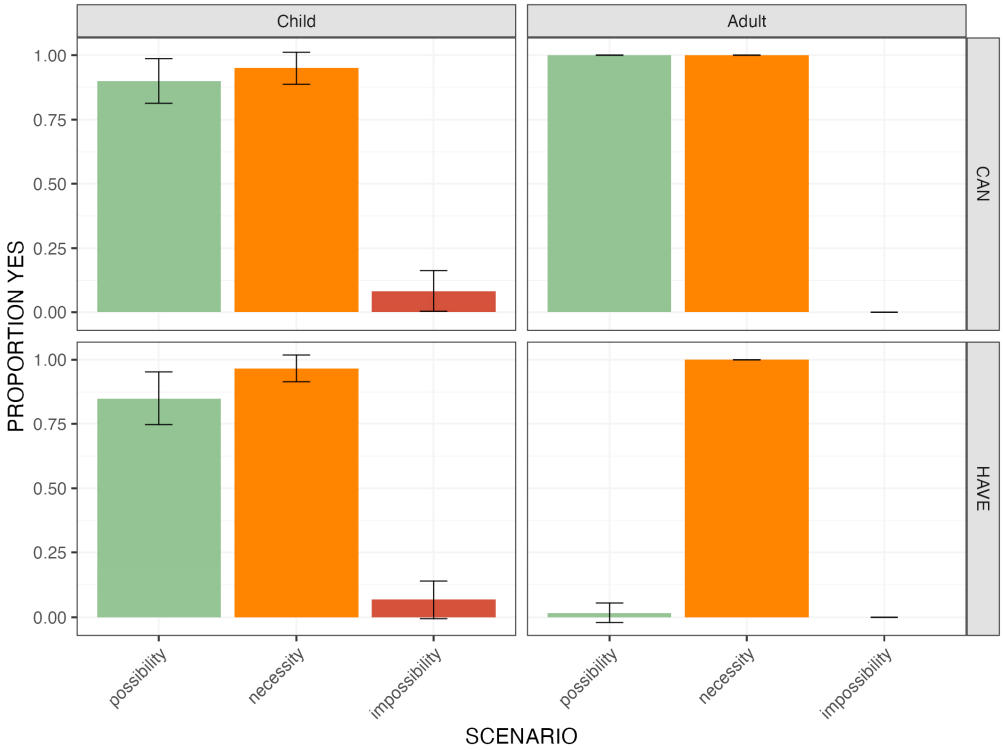
Figure 2
Sample Item From Experiment 1, Showing Flow of Screen Images and Audio



Note. (A) Narrator: "Cat is going to the Balloon Shop to get balloons for the party!." (B) Narrator: "There are two roads to get to the Balloon Shop, the green road and the white road." (C) Narrator: "But uh-oh, the green road is blocked." (D) Logan: "To get to the Balloon Shop, Cat has to go down the white road." Narrator: "Is Logan right?". See the online article for the color version of this figure.

Figure 3

Proportion "Yes" Responses by SCENARIO (POSSIBILITY, NECESSITY, IMPOSSIBILITY)*, Faceted by* AGE *and* MODAL



Note. Adult responses $n = 360$ (180/modal, 60/column), child responses $n = 360$ (180/modal, 60/column). Error bars show errors of proportion. See the online article for the color version of this figure.

IMPOSSIBILITY). The CAN results are along the top, and the HAVE_TO results along the bottom. Adult responses are on the right, child responses are on the left. As illustrated in Table 2, logical responses for *can* sentences are: "yes" in POSSIBILITY, "yes" in NECESSITY, "no" in IMPOSSIBILITY. Logical responses for *have_to* sentences are "no" in POSSIBILITY, "yes" in NECESSITY, "no" in IMPOSSIBILITY. Table 3 provides counts and percentages for each condition.

Adult responses are fully in line with logical responses for both modals. *Can* and *have_to* behave the same way as each other in NECESSITY (orange) and IMPOSSIBILITY (red) scenarios. In POSSIBILITY (green) scenarios, adults vary by modal in the expected way: They always say "yes" with *can*, and almost always say "no" with *have_to*. As in prior studies, adults are judging *can* in possibility scenarios according to its logical meaning. Adult responses confirm that expected logical responses are the same as "adult-like" for this task.

Aggregate child responses for CAN are in line with logical responses, and thus adult-like: In POSSIBILITY and NECESSITY scenarios, children respond "yes" most of the time (90% and 95%, respectively), whereas in IMPOSSIBILITY they give few "yes" responses (8%), as expected. However, for HAVE_TO, child responses are in-line with logical responses in NECESSITY (97% "yes" responses) and IMPOSSIBILITY (7% "yes" responses), but not in POSSIBILITY (85% "yes" 85%). By-individual responses show that 9/19 children are adult-like for CAN ("yes" in possibility and necessity, "no" in impossibility). For HAVE_TO, 8/12 follow the same pattern, with two children showing an adult-like distinction. Remaining children show unexpected patterns (e.g., "yes" in IMPOSSIBILITY).

To assess whether participant responses differed by modal and scenario, we modeled the probability of target logical responses with generalized linear mixed-effect models (Baayen et al., 2008). We used the glmer-function from the "lme4" package available on R to perform our analyses (Bates et al., 2015; R Core Team, 2025). For all models, we estimated model fit (logit link) by the maximum likelihood LaPlace approximation.

The adult data are too invariant to fit, but since their responses are fully in line with expected logical responses, we can indirectly

Table 3

Counts and Percentages of "Yes" Responses in All Conditions, Experiment 1

Age	Modal	POSSIBILITY	NECESSITY	IMPOSSIBILITY
Adults	CAN	60/60 (100%)	60/60 (100%)	0/60 (0%)
	HAVE_TO	1/60 (2%)	60/60 (100%)	1/60 (2%)
Children	CAN	54/60 (90%)	57/60 (95%)	5/60 (8%)
	HAVE_TO	51/60 (85%)	58/60 (97%)	4/60 (7%)

Note. Modals tested between subjects; adults: $n = 2 \times 12$, children: $n = 2 \times 12$. Each participant saw 15 tests items, five per scenario. Values in parenthesis are percentages.

compare child responses to those of adults. We fit a logistic regression model to the child data. The model had target response as the reference for the dependent response variable, independent variables SCENARIO and MODAL were sum-coded, and we included participant and item as random factors, RESPONSE ~ MODAL × SCENARIO + (1|PARTICIPANT) + (1|ITEM), and used the bobyqa optimizer. An additional model was run including age as a continuous variable, but it did not come out as a significant predictor and is therefore left out of the present discussion.

Results of the model reveal main effects for MODAL (less likely to be target with HAVE_TO), SCENARIO (less likely to be target with POSSIBILITY and IMPOSSIBILITY), and for the interaction MODAL × SCENARIO (less likely to be target with HAVE_TO in POSSIBILITY and IMPOSSIBILITY). Post hoc tests were carried out using the emmeans package with the Holms correction for multiple comparisons to further explore this interaction and so all modal-by-scenario pairs could be directly compared (Table 4). These comparisons show a significant difference between the two modals for POSSIBILITY scenarios: Target responses are significantly more likely in POSSIBILITY scenarios for CAN than for HAVE_TO. We also see significant differences occur between HAVE_TO in NECESSITY versus POSSIBILITY (target responses are significantly more likely in NECESSITY), and HAVE_TO in IMPOSSIBILITY versus POSSIBILITY (target responses are significantly more likely in IMPOSSIBILITY). There were no significant differences for CAN versus HAVE_TO in IMPOSSIBILITY or NECESSITY conditions.

Individual response patterns show that 9/12 children on the CAN lists are fully adult-like (counting 4/5 or 5/5 as adult-like for POSSIBILITY and NECESSITY, and 1/5 or 0/5 on IMPOSSIBILITY). For HAVE_TO lists, 8/12 children treat *have_to* exactly the same as *can*. Two children on the HAVE_TO lists are in the adult-like direction, with only 2/5 "yes" in POSSIBILITY, 5/5 in NECESSITY and 0/5 in IMPOSSIBILITY. The remaining two accept *have_to* in IMPOSSIBILITY (but are *can*-like for POSSIBILITY and NECESSITY).

We systematically prompted 144/360 test items for qualitative justifications, and children occasionally gave unprompted justifications, yielding 154 justifications (an additional 21 were excluded; these included pointing and just saying "cuz"). Children on the CAN lists provided 63 responses (POSSIBILITY: 21, NECESSITY: 20, IMPOSSIBILITY: 22). Children on the HAVE_TO lists provided 91 responses (POSSIBILITY: 31, NECESSITY: 28, IMPOSSIBILITY: 32). Explanations in NECESSITY and IMPOSSIBILITY scenarios for both modals nearly always appropriately referenced the blocked road. Focusing on responses in POSSIBILITY scenarios, 19/21(90.5%) explanations for CAN items were with target responses. Of these, seven reference the mentioned road being unblocked, six mentioned that both roads are open, and the remaining eight are uninformative (e.g., "I'm smart," "His pants are yellow"). For HAVE_TO in POSSIBILITY, 6/31 (19.3%) justifications were with target responses. The six justifications for target responses were either uninformative ($n = 3$) or referenced the unmentioned road being open ($n = 3$; 7a). Illustrating the main nonadult pattern of accepting HAVE_TO in POSSIBILITY scenarios, the remaining 25 explanations referred to the two open roads ($n = 9$; 7b,c), repeated the necessity modal or road ($n = 4$; 7d), or were uninformative ($n = 12$).

(7) a. No, the purple

b. Cuz both open

c. Cuz there's yellow road and purple road

d. Cuz he has to

Table 4
Experiment 1: Post Hoc Pairwise Comparisons for Children Comparing Combinations of MODAL by SCENARIO

Comparison	β	*z*	*p*
CAN:POSSIBILITY—HAVE_TO:POSSIBILITY	5.18	5.09	<.0001
HAVE_TO:NECESSITY—HAVE_TO:POSSIBILITY	6.40	6.03	<.0001
HAVE_TO:IMPOSSIBILITY—HAVE_TO:POSSIBILITY	5.61	6.16	<.0001

Experiment 2

Our results from Experiment 1 show that children are adult-like in their responses to possibility *can*, but not *have_to*: their responses to *have_to* are consistent with their interpreting *have_to* as expressing possibility. To test children's comprehension of the modals further, we test them with and without negation in possibility, impossibility, and necessity scenarios.

Method

Participants

For this experiment, we tested two child age groups. The first group was chosen to match the mean age of Experiment 1 (M_{age} = 4;1). The second group was 5-year-olds to see if these children have mastered necessity modals in root modality and to help compare our work to prior studies looking primarily at this age group (Moscati et al., 2017; Ozturk & Papafragou, 2015). Adults were also tested as controls. Because testing was carried out both pre- and post-COVID-19 pandemic, testing for all three age groups was a mixture of both in person and over Zoom (https://zoom.us/), using the secured Zoom license of New York University. Zoom testing was kept as similar to in person as possible, with an experimenter present throughout the session. In all three groups, about three fourths of participants were tested over Zoom. We recruited online child participants using *Children Helping Science* (MIT).

The younger child group participants were 24 monolingual English-speaking children ($n_{female} = 15$, $n_{male} = 9$) ranging from 3;7 to 4;6 (M_{age} = 4;1, SD_{age} = 4 months). An additional 10 children were excluded from the study for not completing or passing training (primarily Zoom participants; passing criteria was the same as for Experiment 1 except now involved choosing both Logan and Lily rather than both "yes" and "no"). The older child group was 24 monolingual English-speaking children ($n_{female} = 13$, $n_{male} = 10$, $n_{unreported} = 1$) ranging in age from 5;0 to 6;0 (M_{age} = 5;7, SD_{age} = 3 months). An additional seven children were excluded for not meeting the experiment demographic requirements. All the older children passed the training. Adult controls were native speakers of English ($n = 24$; $n_{female} = 12$, $n_{male} = 11$, $n_{unreported} = 1$) and ranged from 19 to 34 years (M_{age} = 22, SD_{age} = 3.74).

Sampling criteria and demographic information collected were the same as for Experiment 1. The specific child samples by-modal for the younger child group were *can*: ($n = 12$, $n_{female} = 9$, $n_{male} = 3$; M_{age} = 4;1, SD_{age} = 3.85, $Range_{age}$ = 3;7–4;5) and *have* ($n = 12$, $n_{female} = 6$, $n_{male} = 6$; M_{age} = 4;1, SD_{age} = 3.39, $Range_{age}$ =

3;7–4;6). The specific child samples by-modal for the older child group were *can*: ($n = 12$, $n_{female} = 5$, $n_{male} = 6$, $n_{NA} = 1$; $M_{age} = 5;7$, $SD_{age} = 3.53$, $Range_{age} = 5;0–5;11$) and *have* ($n = 12$, $n_{female} = 8$, $n_{male} = 4$; $M_{age} = 5;7$, $SD_{age} = 3.18$, $Range_{age} = 5;0–6;0$).

Procedure and Experimental Design

This experiment used a sentence-choice task with the same scenarios as in Experiment 1, but presenting both positive and negative statements for CAN (8) and HAVE_TO (9). There were two puppets onscreen, Logan and Lily, who each said one of the modal statements; the participant had to indicate who was right. "Yes" or "no" was added at the beginning of modal statements, because the phonetic distinction between *can* and *can't* in American English is minimal, and to highlight that one puppet always says a positive, the other a negative sentence.

(8) a. Yes, *Cat* **can** *go down the* [mentioned road].

b. No, *Cat* **can't** *go down the* [mentioned road].

(9) a. Yes, *Cat* **has to** *go down the* [mentioned road].

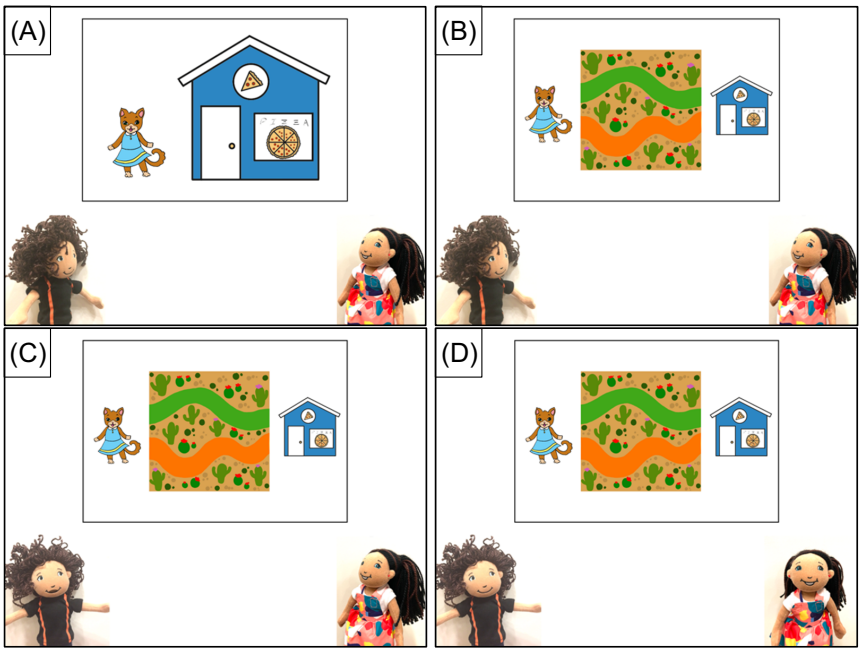
b. No, *Cat* **doesn't have to** *go down the* [mentioned road].

Training was the same as for Experiment 1, except with two puppets instead of one. Logan always appears on the left, Lily on the right, and we counterbalanced which puppet says the positive or negative sentence. Logan always spoke first to aid with memory. Figure 4 illustrates a sample test item in a POSSIBILITY scenario for both modals. The narrator's preamble changed to a question for the puppets (Figure 4B), to pragmatically support either a positive or negative response by the puppet (Figures 4C and 4D). And, the narrator's prompt to the participant changed to, "Who's right? Logan or Lily?" (Figure 4D).

As in Experiment 1, every participant first saw four training items for familiarization to the task. For test items, we varied *MODAL* (CAN, HAVE_TO) between subjects, and *SCENARIO* (POSSIBILITY, NECESSITY, IMPOSSIBILITY) within subjects. Half the participants were assigned to CAN lists ($n = 12$ adults, 12 3.5–4.5yos, 12 5yos), and half were assigned to HAVE_TO lists ($n = 12$ adults, 12 3.5–4.5yos, 12 5yos). There were four different lists for each modal, for a total of eight lists. Three participants from each age group were run on each list. Participants all saw 16 test items (order randomized and varied by list) with eight items with POSSIBILITY scenarios and four each with NECESSITY and IMPOSSIBILITY scenarios. We opted to have more POSSIBILITY scenarios following our prior results, as this is the scenario of greatest interest. Order of mention of roads, mentioned color, and target road position was fully counterbalanced. The dependent variable was which sentence/puppet the participant chose (negated or positive modal). We prompted about one half of the items (marked on scoresheets with an asterisk) with "How come?" for qualitative responses. The experimenter recorded all participant responses by hand on a scoresheet. Most sessions were recorded (when parental consent was obtained), allowing double-checking of responses.

Figure 4

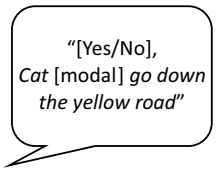
Sample Item From Experiment 2, Showing Flow of Screen Images and Audio



Note. (A) Narrator: "Cat is going to the Pizzeria to get a pizza for the party". (B) Narrator: "There are two ways to get to the Pizzeria, the green road and the orange road. To get to the Pizzeria, *can* Cat [does Cat *have to*] go down the orange road?". (C) Logan: "Yes, Cat *can* [*has to*] go down the orange road." (D) Lily: "No, Cat *can't* [*doesn't have to*] go down the orange road." Narrator: "Who's right? Logan or Lily?". See the online article for the color version of this figure.

Table 5

Logical Responses (Can vs. Can't; Have_To vs. Doesn't Have_To) by SCENARIO *and* MODAL



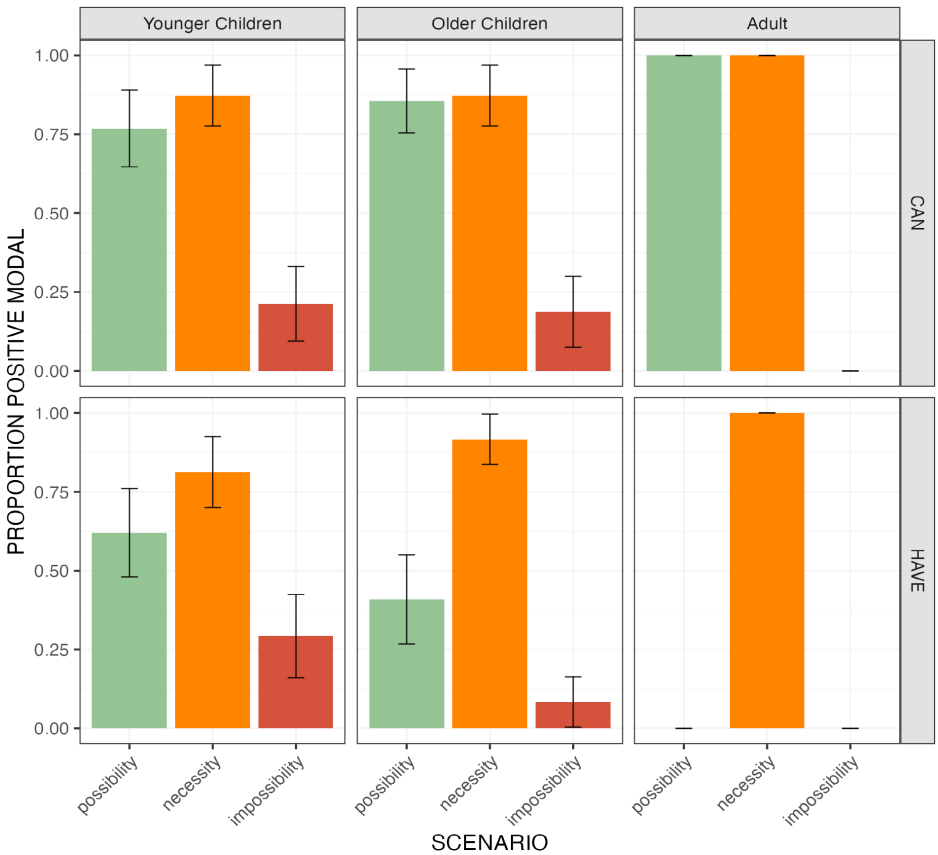
		Scenario		
		POSSIBILITY	NECESSITY	IMPOSSIBILITY
				
	Modal			
Logical Response	CAN	can	can	can't
	HAVE_TO	doesn't have_to	has_to	doesn't have_to

Note. See the online article for the color version of this table.

Modal claims and their negated counterparts have opposite truth values, so for each scenario, one puppet always says something logically true and the other something logically false (*can* vs. *can't*; *have_to* vs. *doesn't have_to*). The expected logically true choices for the modals by scenario are provided in Table 5. We provide the positive or negated modal to illustrate the key contrasts. In NECESSITY scenarios, we expect (adult) participants to pick *can* over *can't*, and *have_to* over *doesn't have_to*. In IMPOSSIBILITY scenarios, we expect them, conversely, to pick *can't* over *can*, and *doesn't have_to* over *have_to*. In POSSIBILITY scenarios, crucially, we expect them to pick *can* over *can't*, but *doesn't have_to* over *have_to*. As in Experiment 1, the two modals pattern alike in NECESSITY and IMPOSSIBILITY but diverge in POSSIBILITY.

Figure 5

Proportion of Positive Modal Responses by SCENARIO*, Faceted by* AGE *and* MODAL



Note. Adult responses: $n = 384$ (192/modal), younger children responses: $n = 372$ (189 CAN, 183 HAVE_TO), older children responses: $n = 375$ (185 CAN, 189 HAVE_TO). Error bars show errors of proportion. See the online article for the color version of this figure.

Results

Figure 5 shows target responses to the prompt "Who's right? Logan, or Lily?" calculated from puppet choice based on whether the puppet said the positive (*can/have_to*) or negated (*can't/doesn't have_to*) modal. The figure is organized by modal (CAN, HAVE_TO) and which scenario was depicted (POSSIBILITY, NECESSITY, IMPOSSIBILITY). Responses are arranged from youngest to oldest: Child facets are left (younger) and middle (older), adult facets at the right. The *x*-axis corresponds to the proportion of positive responses, that is, when the participant chose the puppet who said the positive modal statement (*can*, *have_to*). Table 6 provides counts and percentages by condition.

Adult responses are fully in line with expected responses for all modals and scenarios. For CAN, they choose *can* in POSSIBILITY and NECESSITY, and *can't* in IMPOSSIBILITY. For HAVE_TO, they choose *have_to* in NECESSITY, and *doesn't have_to* in POSSIBILITY and IMPOSSIBILITY. These results pattern with Experiment 1 but with a sentence-choice task and negation.

Younger children's responses for CAN are in the direction of adult responses for all three scenarios, but with many nonadult responses across the board. We excluded nonresponses from the analyses (i.e., cases where the child did not give an answer): One response was excluded in each situation condition. Of the remaining, in POSSIBILITY scenarios, 77% of responses were adult-like selections of *can*. In NECESSITY scenarios, 87% were adult-like selections of *can*. In IMPOSSIBILITY scenarios, 79% of responses were selections of adult-like negated *can't*. Two younger child participants were fully adult-like for CAN.

Younger children's responses for HAVE_TO do not show the adult-like pattern, with nonadult responses in all scenarios, especially POSSIBILITY. In POSSIBILITY scenarios, we excluded one nonresponse, and eight "both" responses (all from the same child). We decided prior to testing to exclude "both" responses if they obtained, rather than treat them as nontarget, because they do not conform to our task or statistical models. Of the remaining, 62% were nonadult-like positive *have_to*. A one-tailed binomial test[3] reveals this result is significantly different from chance in the unexpected (nonadult-like) direction (54/87, $p = .02$). The nonadult positive modal responses came from 11/12 participants (N.B. the remaining participant was also nonadult-like due to all "both" responses). In NECESSITY scenarios, 81% of responses were positive modal *have_to*. In IMPOSSIBILITY scenarios, 71% of responses were adult-like negated *doesn't have_to*. No younger children were fully adult-like for HAVE_TO.

Table 6
Counts and Percentages of Positive Modal Responses by Age for All Conditions, Experiment 2

Age	Modal	POSSIBILITY	NECESSITY	IMPOSSIBILITY
Adults	CAN	96/96 (100%)	48/48 (100%)	0/48 (0%)
	HAVE_TO	0/96 (0%)	48/48 (100%)	0/48 (0%)
Younger children	CAN	73/95 (77%)	41/47 (87%)	10/47 (21%)
	HAVE_TO	54/87 (62%)	39/48 (81%)	14/48 (29%)
Older Children	CAN	77/90 (86%)	41/47 (87%)	9/48 (19%)
	HAVE_TO	38/93 (41%)	44/48 (92%)	4/48 (8%)

Note. Modals tested between subjects; adults: $n = 2 \times 12$, children: $n = 2 \times 12$. Each participant saw 16 test items, eight items for POSSIBILITY, four each for NECESSITY and IMPOSSIBILITY. Values in parenthesis are percentages.

Older children's responses for CAN are in the direction of adult responses for all three scenarios and resemble the younger children's responses for CAN with many nonadult responses across the board. We excluded some trial responses from the final analysis, these included six "both" responses from a single participant in POSSIBILITY scenarios, and one "both" in NECESSITY scenarios. Of the remaining responses, 86% were adult-like selections of *can* in POSSIBILITY, 87% were adult-like selections of *can* in NECESSITY, and 19% were nonadult-like selections of *can* in IMPOSSIBILITY. Three participants were fully adult-like for CAN.

Older children's responses for HAVE_TO look somewhat more adult-like than those of the younger children, but still show nonadult responses in all scenarios, especially POSSIBILITY. In POSSIBILITY scenarios, we excluded three "both" response trials. Of the remaining, 41% were nonadult-like positive *have_to*. A one-tailed binomial test reveals that this result is significantly different from chance in the expected (logical, adult-like) direction (38/93, $p = .05$). In NECESSITY scenarios, 92% of responses were positive *have_to*. In IMPOSSIBILITY scenarios, 8% of responses were nonadult positive *have_to*. One older child was fully adult-like for HAVE_TO.

We modeled target responses using the same R packages and assumptions as for Experiment 1. To assess whether participant responses of positive or negated sentence choice differed by modal and by scenario, we fit two logistic regression models to the child data, one for the younger children, one for the older children. The models had target response as a reference for the dependent variable (this is functionally equivalent to adult-like, as adults performed perfectly in-line with logical predictions), sum-coding for the independent variables for SCENARIO and MODAL and included participant and item as random factors, TARGET ~ MODAL × SCENARIO + (1|PARTICIPANT) + (1|ITEM), and used the bobyqa optimizer. We also ran a model including age as a continuous variable but found no significant effect of age ($p = .88$).

Results of the model for the younger children reveal main effects for MODAL (less likely to be target with HAVE_TO) and SCENARIO (less likely to be target in POSSIBILITY). Post hoc tests were carried out using the emmeans package with the Holms correction for multiple comparisons to further explore this interaction and so all modal-by-scenario pairs could be directly compared (Table 7). These comparisons show a significant difference between the two modals for POSSIBILITY scenarios: Target responses are significantly more likely for CAN than HAVE_TO. We also see significant differences between HAVE_TO in NECESSITY and POSSIBILITY (target responses more likely in NECESSITY), and HAVE_TO in IMPOSSIBILITY and POSSIBILITY (target responses more likely in IMPOSSIBILITY). There were no significant differences for CAN versus HAVE_TO in IMPOSSIBILITY or NECESSITY conditions.

Results of the model for the older children reveal no significant main effects. Post hoc tests were carried out using the emmeans package with the Holms correction for multiple comparisons (Table 8). These comparisons show significant differences between HAVE_TO in NECESSITY and POSSIBILITY (target responses more likely in NECESSITY), and HAVE_TO in IMPOSSIBILITY and POSSIBILITY (target responses more likely in IMPOSSIBILITY). There were no significant differences for CAN versus HAVE_TO in IMPOSSIBILITY, as with the younger children. The difference between modals for POSSIBILITY

[3] We conducted one-tailed tests because we had a predicted direction of effect, and it provides more power.

Table 7
Experiment 2: Post Hoc Pairwise Comparisons for Younger Children Comparing Combinations of MODAL *by* SCENARIO

Comparison	β	*z*	*p*
CAN:POSSIBILITY—HAVE_TO:POSSIBILITY	1.90	4.31	.0002
HAVE:NECESSITY—HAVE_TO:POSSIBILITY	2.19	4.40	.0002
HAVE:IMPOSSIBILITY—HAVE_TO:POSSIBILITY	1.56	3.44	.0077

scenarios is significant for the older children (target responses more likely for CAN than for HAVE_TO in possibility.)

To get more information on the development of HAVE_TO comprehension, we also ran post hoc tests on all the child data (treating age group categorically with two levels), in order to run emmeans to directly compare whether older children were significantly more adult-like than younger children for HAVE_TO, in particular in POSSIBILITY. The model was the same as the others, adding AGE as a factor (sum coded). Results for all children reveal significant main effects for SCENARIO (less likely to be target in POSSIBILITY), AGE (more likely to be target if OLDER), and for MODAL × SCENARIO (less likely to be target with HAVE_TO in IMPOSSIBILITY). Post hoc tests were carried out using the emmeans package with the Holms correction for multiple comparisons. We were specifically concerned with the two comparisons in Table 9. These show that OLDER children are significantly more adult-like for HAVE_TO overall than younger children, but that they do not significantly differ from younger children for HAVE_TO in POSSIBILITY.

Children's individual response patterns show considerable variability, more so than in Experiment 1. We include by-individual plots in our Open Science Framework repository and histograms in Figure 6.[4]

These histograms clearly show (i) noise across the board, (ii) particular difficulty with POSSIBILITY, and (iii) more difficulty with *have_to* than *can*. In reference to Figure 6, and the by-individual plots, we highlight a few important points here. First, across both age groups and both modals, about one third of children selected the positive modal (*can*, *have_to*) in *impossibility* scenarios to some degree—a surprising pattern that may reflect task demands on working memory. For *can*, most children in both groups show broadly adult-like patterns, and a few children show more haphazard patterns. For *have_to*, most younger children skew toward the pattern whereby *have_to* is nonadult-like and patterns similarly to *can* (with more nontarget responses in POSSIBILITY). Among the younger group, four children are in the target direction for all three scenarios. Among the older group, one child is fully adult-like for *have_to*, and four others are in the adult-like direction. However, for responses for *have_to* in POSSIBILITY many older children show a mix of target and nontarget responses, rather than consistent patterns aligned with interpreting *have_to* as either possibility or necessity. Thus, while younger children skew nonadult-like for *have_to*, older children appear more mixed—some trending adult-like, others showing at-chance patterns.

Table 8
Experiment 2: Post Hoc Pairwise Comparisons for Older Children Comparing Combinations of MODAL *by* SCENARIO

Comparison	β	*z*	*p*
HAVE_TO:NECESSITY—HAVE_TO:POSSIBILITY	2.13	3.72	.0030
HAVE_TO:IMPOSSIBILITY—HAVE_TO:POSSIBILITY	2.13	3.72	.0030
CAN:POSSIBILITY—HAVE_TO:POSSIBILITY	1.53	3.37	.0098

We systematically prompted half of the test items (384/768) for qualitative justifications, and children also provided some unprompted ones. This experiment collected more data than Experiment 1, and due to the puppet-choice task, responses often defaulted to "because she's right." Therefore, we focus on *have_to* in POSSIBILITY. As in Experiment 1, explanations in NECESSITY and IMPOSSIBILITY scenarios typically referenced the blocked road, with most aligning with target responses. However, qualitative responses for *have_to* in the POSSIBILITY indicate that younger children were not adult-like, 32 explanations, 12 (38%) with target responses. Those with nonadult responses (positive) sometimes used possibility modals to describe open paths ($n = 4$; 10a), repeated *have_to* or another necessity modal ($n = 3$; 10b), or justified their agreement with the *have_to* puppet ($n = 3$). Other responses were inconsistent or unclear. Those with adult-like responses (negative) sometimes repeated *have_to* ($n = 5$), and one gave compelling adult-like responses (10c). Qualitative responses from 5-year-olds were more mixed, $n = 54$; 34 (63%) with target answers. For nonadult responses (positive), children sometimes used a possibility modal to reword what they had heard (*could, can*; $n = 4$; 11a), repeated *have_to* or another necessity modal ($n = 3$), or found a way to justify why they picked the puppet ($n = 9$; 11b). For adult-like *doesn't have_to* responses, children sometimes gave very cogent answers, showing adult-like comprehension and often referring to open possibilities with a possibility modal or negated necessity ($n = 10$; 11c).

(10) a. Can go down the pink one and the purple one

b. Because she does need to go down the pink road

c. Because cat can go there [unmentioned road]

(11) a. Cat can go down the yellow road and she can also go down the white road

b. Because there's no stoppage and that's the easier way to get there

c. The cat doesn't have to go down the white road he can take the yellow road too

Discussion

In both experiments, adult results show the logically expected pattern whereby *can* and *have_to* pattern the same way in NECESSITY and IMPOSSIBILITY scenarios, but opposite from each other in POSSIBILITY. Focusing first on Experiment 1, the child results go against the PREMATURE CLOSURE HYPOTHESIS. Premature closure predicts that children accept necessity modal *have_to* in possibility situations about half the time. However, children were significantly below chance for *have_to* in possibility scenarios, but adult-like for *have_to* in NECESSITY and IMPOSSIBILITY, and for *can* in all scenarios

[4] We thank an anonymous reviewer for recognizing and emphasizing the contribution of this data analysis.

Table 9

Experiment 2: Post Hoc Pairwise Comparisons for Older Versus Younger Children for HAVE_TO in General and in POSSIBILITY

Comparison	β	*z*	*p*
HAVE_TO:YOUNGER—HAVE_TO:OLDER	−1.15	−2.87	.0215
HAVE_TO:POSSIBILITY:YOUNGER—HAVE_TO:POSSIBILITY:OLDER	−0.96	−2.37	.4301

(which also contradicts premature closure). No condition was at chance. Children treated *have_to* the same as *can* in all three scenarios. Results are consistent with children (M_{age} = 4;1) not having adult-like force for *have_to* (NONADULT FORCE HYPOTHESIS), as they accept it in possibility scenarios where it is false for adults. Importantly, using teleological modals rather than epistemic did not improve children's performance compared to prior studies. Rather, teleological modals and contexts reveal that reasoning about open possibilities is likely more robust by age 4 than often assumed, but word-knowledge—particularly for necessity modals—is not yet adult-like.

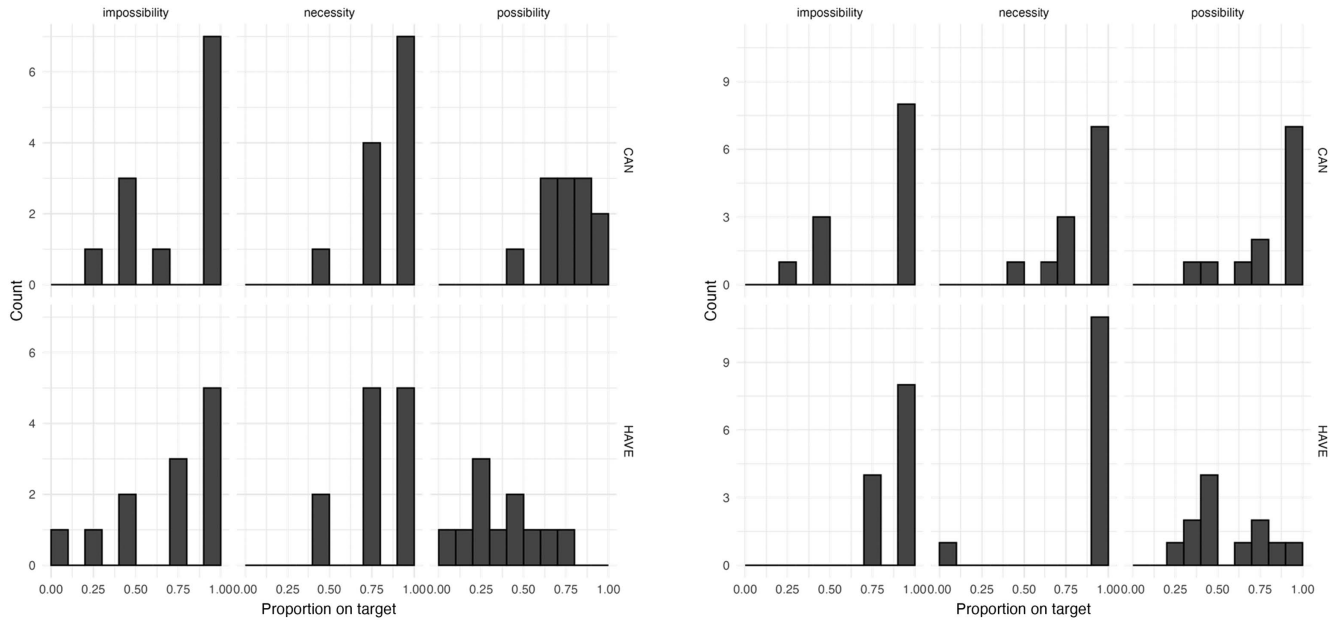
Our child sample in Experiment 1 is about a year younger than prior studies that argue for premature closure (Moscati et al., 2017; Ozturk & Papafragou, 2015). It would be surprising if children no longer showed developmental difficulties with representing and maintaining open possibilities at age 4, but began to re-exhibit those difficulties at age 5. More consistent with our results is that children are able to represent open possibilities by age 4, consistent with nonlinguistic tasks like the forked tube task (e.g., B. P. Leahy & Carey, 2020; Redshaw & Suddendorf, 2016), but are still learning the force of their input modals, in particular necessity modals. Prior studies with 5-year-olds likely underestimated children's conceptual abilities, due to testing in epistemic modality, which introduces potential task effects with perspective-taking, and relying on rarely used modals (like *may*) or modals that are overwhelmingly used in nonepistemic flavors in the input (like *have_to*; van Dooren et al., 2022). Conversely, these studies likely overestimated children's linguistic knowledge, as they assumed knowledge of both force and flavor, despite modal words presenting learners with a particularly complex mapping challenge.

Child results from Experiment 2 show more noise than those of Experiment 1, for both child age groups. We attribute this to the processing complexity increase for this task—children have to listen to two sentences, and parse and remember both modality and negation, and who said what. Despite this, our results for younger children (M_{age} = 4;1) show that *can* is adult-like in all scenarios, while *have_to* is adult-like in NECESSITY and IMPOSSIBILITY, but not POSSIBILITY (as in Experiment 1). In possibility, *have_to* results look less sharp than Experiment 1—but they remain not at chance, and the effect is in the nonadult direction. We conclude that results from Experiment 2 for younger children seem to go against the PREMATURE CLOSURE HYPOTHESIS: Neither possibility condition is at chance. But, neither are these results fully consistent with the NONADULT FORCE HYPOTHESIS that child *have_to* has a possibility semantics, meaning possible in the positive, and impossible when negated; this pattern arises in the aggregate (albeit not as cleanly as in Experiment 1), but the NONADULT FORCE HYPOTHESIS implies that individuals should uniformly treat *have_to* as a possibility (or as a necessity if they have become adult-like)—some children in this sample fit this pattern, albeit with some noise, but others show more random behaviors that are not consistent with having a possibility semantics for *have_to* (nor a necessity semantics).

For older children, results also allow us to reject the PREMATURE CLOSURE HYPOTHESIS, because *can* is largely adult-like in possibility

Figure 6

Experiment 2: Histogram Plots for the Younger Children (Left) and Older Children (Right), Showing Counts of Individuals by Proportion of Responses for Each Modal and Scenario



scenarios, and *have_to* is slightly better than chance in the adult-like direction. But the NONADULT FORCE HYPOTHESIS is not supported. The aggregate statistics are consistent with the NONADULT FORCE HYPOTHESIS predictions, as older children are overall less adult-like for *have_to* (over *can*), less adult-like for *have_to* in POSSIBILITY (over NECESSITY or IMPOSSIBILITY), and more adult-like for *can* than *have_to* in POSSIBILITY. But the distribution of individual 5-year-olds' responses is not as expected under the NONADULT FORCE HYPOTHESIS (with children treating *have_to* as possibility) but rather shows more random behavior. A handful of 5-year-olds appear to be behaving like adults for *have_to*, but this age group overall shows a pattern inconsistent with both hypotheses: largely adult-like for *can* and *have_to* in necessity and impossibility, adult-like for *can* in possibility, but at-chance for *have_to* in POSSIBILITY.

Children's variable response patterns in Experiment 2 may reflect additive effects of online processing demands. First, children showed general difficulty with the task—evidenced by a high number of incorrect responses in IMPOSSIBILITY scenarios, likely due to working memory constraints. Second, nonadult-like responses were more common with *have_to* than *can*, as in Experiment 1. Third, children showed the largest deviations from adult-like behavior in POSSIBILITY scenarios, likely because reasoning through multiple open possibilities is especially cognitively demanding. As we saw in Experiment 1, in a simple Truth Value Judgement Task, young children can deal with reasoning through multiple possibilities, but the added cognitive demands of Experiment 2—having to remember and choose between two statements—may have been too demanding. The additive difficulty of the modal (*have_to*) and the scenario (POSSIBILITY) may therefore underlie particularly noisy performance in the *have_to*—POSSIBILITY condition. Experiment 2 was initially designed as a TVJT (see our Open Science Framework for details), but children exhibited systematic difficulties in only those conditions involving true negative sentences or false positive sentences, leading us to revise the task. Future research should aim to reduce task demands to better isolate negated modal comprehension and explore the within-subjects relationship between working memory and modal comprehension.

Experiment 2 results show a pattern largely consistent with *have_to* being possibility, and *doesn't have_to* impossibility for younger children; these results are also consistent with these children thinking *doesn't have_to* is a nonadult "necessary not" (as in prior studies testing modal-negation scope), but then children would be treating *have_to* as a possibility, except when negated. The unified analysis, where *have_to* is possibility and *doesn't have_to* is impossibility (just like *can't*) seems, a priori, more likely, and consistent with previous findings that children tend to prefer surface scope with scope bearing elements like negation (see Lidz & Musolino, 2002). Underlying force issues may also explain why issues with modal scope comprehension (taken instead as both issues with basic semantic force *and* interactions with negation) persist longer than with other quantifiers (Moscati et al., 2016).

Another possible explanation for our results preserves the PREMATURE CLOSURE HYPOTHESIS, by adding a pragmatic *mention-bias*.[5] In our tasks, children might drop the unmentioned option, assuming the speaker's mention is meaningful. This could account for our Experiment 1 results (though in qualitative responses, both options are often mentioned). This explanation is orthogonal to force interpretation: If children discard the unmentioned option in possibility scenarios, we cannot assess whether they interpret *can* and *have_to* correctly. It also challenges the typical assumption that children interpret the modal sentence *after* discarding one option—an assumption needed to explain chance-level modal judgments with premature closure (Bleotu et al., 2021; Moscati et al., 2017; Ozturk & Papafragou, 2015). Elicited modal production data from Cournane, Hirzel, and Hacquard (2023) indicate that children at this age readily produce *can* and *have_to*, but they often fail to distinguish possibility modals (*can, could, might*) from necessity modals (*have_to, must*), using them in the same situations. Although a mention-bias may contribute, these findings more strongly support the interpretation that children have not yet acquired a distinction in modal force. Moreover, in our experiments, children consistently repeat the modal in their qualitative responses. Taken together, these results also confirm that children are attending to and responding to the modal in our tasks, rather than disregarding it.

Despite our efforts to promote teleological readings and suppress deontic interpretations, English modals are flexible in flavor—including subtypes within root modality—and there is always a risk that participants interpret them in unintended ways. A small number of children's qualitative responses are consistent with a deontic reading (interpreting *have_to* as an order from the puppet), though never unequivocally. However, such cases are infrequent, and overall the stimuli and results do not strongly support a deontic explanation. Recall, the puppet was portrayed as a peer—a child—who spoke in a deliberately dopey and nonauthoritative tone. And, treating the puppet as an authority in IMPOSSIBILITY scenarios would conflict with contextual cues (and most children did well in these scenarios). Because scenarios were randomized, such cases would actively counteract any persistent authority bias. Notably, even if children did access a deontic interpretation at times, this would reflect challenges with form-meaning mapping, rather than conceptual difficulty with possibility reasoning. In the future, materials could be normed for flavor, to the level of specific subflavors (i.e., deontic vs. teleological, not just root vs. epistemic) to try to further increase confidence that participants access the intended flavor.

Another nonadult semantics for child modals proposed in the literature is Staniszewski et al. (2023), which takes children's tendency to accept possibility modals in necessity contexts to result from a default assumption that all modals are "homogeneous." Homogenous modals are modals like *should*, which differ from regular possibility or necessity modals in patterns of contradictions in conjunctions: statements of the form "*modal* p and *modal* not p" yield contradictions with necessity modals, but not possibility modals, statements of the form "not *modal* p and not *modal* q" yield contradictions with possibility, but not necessity modals, while homogeneous modals yield contradictions with both. This analysis for child modals is inconsistent with our results as, if children's *can* and *have_to* were homogeneous modals, we would expect children to reject them in possibility situations, contrary to fact.

Future research should include a broader range of child populations, languages, and modal verbs to better assess generalizability, as our current samples are relatively small and skew toward children of highly educated parents. Using the same modal verbs as B. Leahy and Zalneiriunas (2021) in future-oriented modal situations, we observed similar response patterns, strengthening confidence that our Experiment 1 results reflect younger preschoolers' tendency to

[5] We thank an anonymous reviewer for this suggestion.

treat *have_to* similarly to *can*. Further work is needed to clarify distinctions among subflavors of root modality and to compare root with epistemic modality. In contrast to our and Leahy and Zalneiriunas' results, studies using unambiguous epistemic modals in present-oriented contexts report persistent, nonadult-like, at-chance performance with both possibility and necessity modals (Bleotu et al., 2021; Moscati et al., 2017). These results highlight the need for within-participant comparisons of root and epistemic modals to isolate flavor effects on children's modal language and reasoning.

In sum, our results suggest preschool-aged children's struggles with modal force may not be primarily caused by issues with reasoning over open possibilities or nonadult scope biases. Rather, they could stem from a nonadult force mapping, possibly related to the subset problem that modals pose: necessity entails possibility, hence uses of necessity modals are consistent with a weaker possibility meaning. Our data suggest that children have a possibility meaning bias for modals (cf. Crain et al., 1994). Why should children assume a weak semantics for strong input modals, and how do they eventually arrive on the correct meaning? We speculate that learners might assume that a dedicated modal word may be more likely to express possibility, as necessities can be conveyed through other means, for example, with bare assertions for epistemic necessity ("Nick is in the red box;" noted in Noveck, 2001, and Ozturk & Papafragou, 2015), and with imperatives for root necessity ("Take the red road!" vs. "You have to take the red road").

As to how children eventually figure out the correct force of necessity modals, we appeal to children's input experience. They may get some cues for force from how modals interact with negation (see Gualmini & Schwarz, 2009), but necessity modals vary in how they interact with negation (i.e., *have_to* scopes under, but *must* scopes above) adding complications for this route (Iatridou & Zeijlstra, 2013). Perhaps more promising, the pragmatic and situational contexts in which caregivers use modals may be a rich learning source (for data and discussion, see Dieuleveut et al., 2022; for a similar case study see Rasin & Aravind, 2021). Dieuleveut et al. (2022) found promising results that the desirability of the modal prejacent may help pragmatically bootstrap the force of root modals: We typically *have_to* do undesirable things like brush our teeth but *can* do desirable things like play games. Future research should continue to investigate how necessity modals become adult-like. Cournane, Dieuleveut, and Hacquard (2023) looked at correlation data for the mother–child input–output dyads discussed in Dieuleveut et al. (2022) and found two significant input factors predicting how adult-like 2-year-olds' necessity modals were: the amount of modal talk in their input in general and specifically the amount of negated necessity modals.

Our results do not rule out B. Leahy and Zalneiriunas's (2021) proposal, which could also account for our data patterns (particularly for the younger child samples). On their view, children in our experiments may rely on minimal representations of possibility, focusing only on the mentioned-road possibility because they reason about one possibility at a time. While this remains a plausible explanation, we favor our own interpretation. The idea that young children lack modal representations of possibility (B. P. Leahy & Carey, 2020) becomes increasingly difficult to maintain as children get older—especially given that most comprehension studies on modal force test children over 4. We also find that children's qualitative responses often mention both roads, even when their quantitative patterns are nonadult. Our account aligns with the continuity hypothesis (Pinker, 1984). Although this is only a theoretical preference, we view the minimal representations of possibility proposal as introducing two layers of nonadultness: a conceptual discontinuity which must be overcome, and a separate word-learning challenge as children would still need to learn the meanings of individual modal terms. Moreover, the account currently lacks a clear mechanism for how linguistic experience might unlock the development of logical reasoning (as B. P. Leahy & Carey, 2020 propose). In this sense, we believe our account may offer a more parsimonious explanation. That said, both perspectives raise valuable questions, and we see merit in future work that directly compares them, ideally with 3-year-olds.

Conclusion

This article presented two experiments testing children's understanding of the force of modals *can* and *have_to* in teleological (goal-oriented) scenarios. Our goal was to test whether children understand the basic force semantics of these modals, namely that *can* expresses possibility, and *have_to* necessity. Prior studies have shown that children tend to accept necessity modals (and sometimes possibility modals) in possibility situations about half of the time, and they have taken this behavior to stem from conceptual difficulties reasoning about open possibilities: in the face of indeterminacy, children perform "premature closure," randomly throwing out some possibilities. Our results show instead that preschoolers seem to have an adult-like understanding of possibility *can*, but through age 4, and perhaps for some time after, they treat necessity *have_to* as a possibility modal, in affirmative and arguably in negative sentences for younger children. By age 5, some children are beginning to behave like adults, but further research and improved methods are required to tease apart linguistic and processing factors, especially in negative sentences. Our results suggest that preschool-aged children's difficulty with modal force tasks may be due to word learning challenges, especially discovering which modals express necessity.

References

Acredolo, C., & Horobin, K. (1987). The child's relational reasoning and tendency toward premature closure. *Developmental Psychology*, *23*(1), 13–21. https://doi.org/10.1037/0012-1649.23.1.13

Alderete, S., & Xu, F. (2023). Three-year-old children's reasoning about possibilities. *Cognition*, *237*, Article 105472. https://doi.org/10.1016/j.cognition.2023.105472

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, *60*(1), 40–62. https://doi.org/10.1016/j.cogpsych.2009.06.002

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*(1), 84–93. https://doi.org/10.1016/j.cognition.2010.10.010

Bascelli, E., & Barbieri, M. S. (2002). Italian children's understanding of the epistemic and deontic modal verbs dovere (must) and potere (may). *Journal of Child Language*, *29*(1), 87–107. https://doi.org/10.1017/S0305000901004925

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beck, S. R., & Robinson, E. J. (2001). Children's ability to make tentative interpretations of ambiguous messages. *Journal of Experimental Child Psychology*, *79*(1), 95–114. https://doi.org/10.1006/jecp.2000.2583

Bleotu, A. C., Benz, A., & Gotzner, N. (2021). Shadow playing with Romanian 5-year-olds. Epistemic adverbs are a kind of magic! *Experiments in Linguistic Meaning*, *1*, 59–70. https://doi.org/10.3765/elm.1.4866

Bleotu, A. C., Slăvuțeanu, G., & Benz, A. (2023). The role of intonation and context in lack of necessity meanings in negated deontic necessity modals in child Romanian. *Proceedings of semantics and linguistic theory* (pp. 541–562). Semantics and Linguistic Theory.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. MIT Press.

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* (Version 5.3.51) [Computer program]. https://www.fon.hum.uva.nl/praat/

Byrnes, J. P., & Duff, M. A. (1989). Young children's comprehension of modal expressions. *Cognitive Development*, *4*(4), 369–387. https://doi.org/10.1016/S0885-2014(89)90049-X

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science 359*, 1263–1266. https://doi.org/10.1126/science.aao3539

Cesana-Arlotti, N., Varga, B., & Téglás, E. (2022). The pupillometry of the possible: An investigation of infants' representation of mutually exclusive alternatives. *Philosophical Transactions of the Royal Society B*, *377*(1866), Article 20210343. https://doi.org/10.1098/rstb.2021.0343

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H.-J. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of the 25th annual Boston University conference on language development* (Vol. 25, pp. 157–168). Cascadilla Press.

Chierchia, G., Guasti, M. T., Gualmini, A., Meroni, L., Crain, S., & Foppolo, F. (2004). Semantic and pragmatic competence in children's and adults' comprehension of *Or*. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 283–300). Palgrave Macmillan. https://doi.org/10.1057/9780230524125_13

Choi, S. (1988). The semantic development of negation: A cross-linguistic longitudinal study. *Journal of Child Language*, *15*, 517–531. https://doi.org/10.1017/S030500090001254X

Cournane, A. (2021). Revisiting the epistemic gap: It's not the thought that counts. *Language Acquisition*, *28*(3), 215–240. https://doi.org/10.1080/10489223.2020.1860054

Cournane, A., Dieuleveut, A., & Hacquard, V. (2023). Assessing how the linguistic input affects children's mastery of modals. *Proceedings of the 47th Annual Boston University Conference on Language Development (BUCLD47)* (pp. 173–186). Cascadilla Press.

Cournane, A., Hirzel, M., & Hacquard, V. (2023). Mapping modal verbs to meanings: An elicited production study on "force" and "flavor" with young preschoolers. *Language Acquisition*, *31*(1), 57–80. https://doi.org/10.1080/10489223.2023.2173074

Cournane, A., & Tailleur, S. (2020). La production épistémique chez l'enfant francophone: Complexité syntaxique et ordre d'acquisition [Epistemic productions in French-speaking children: Syntactic complexity and order of acquisition]. *Arborescences: Revue D'études Françaises*, *10*, 47–72. https://doi.org/10.7202/1081888ar

Crain, S. (2012). *The emergence of meaning*. Cambridge University Press. https://doi.org/10.1017/CBO9780511842863

Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 443–468). Lawrence Erlbaum.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar*. MIT Press.

Cummins, D. D. (1996). Evidence of deontic reasoning in 3- and 4-year-old children. *Memory & Cognition*, *24*(6), 823–829. https://doi.org/10.3758/BF03201105

Dack, L. A., & Astington, J. W. (2011). Deontic and epistemic reasoning in children. *Journal of Experimental Child Psychology*, *110*(1), 94–114. https://doi.org/10.1016/j.jecp.2011.04.003

de Carvalho, A., Crimon, C., Barrault, A., Trueswell, J., & Christophe, A. (2021). "Look! It is not a bamoule!": 18- and 24-month-olds can use negative sentences to constrain their interpretation of novel word meanings. *Developmental Science*, *24*(4), Article e13085. https://doi.org/10.1111/desc.13085

de Carvalho, A., Gomes, V., & Trueswell, J. (2023). English-learning preschoolers can correctly parse and interpret negative sentences to guide their interpretations of novel noun and verb meanings. *Language Acquisition*, *30*(3–4), 277–310. https://doi.org/10.1080/10489223.2022.2137805

Deal, A. R. (2011). Modals without scales. *Language*, *87*(3), 559–585. https://doi.org/10.1353/lan.2011.0060

Dieuleveut, A. (2021). *Finding modal force* [Doctoral dissertation, University of Maryland, College Park]. ProQuest Dissertations Publishing.

Dieuleveut, A., Cournane, A., & Hacquard, V. (2021). Finding the force: A novel word learning experiment with modals. *Experiments in Linguistic Meaning*, *1*, 136–146. https://doi.org/10.3765/elm.1.4883

Dieuleveut, A., van Dooren, A., Cournane, A., & Hacquard, V. (2022). Finding the force: How children discern possibility and necessity modals. *Natural Language Semantics*, *30*(3), 269–310. https://doi.org/10.1007/s11050-022-09196-4

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). Academic Press. https://doi.org/10.1163/9789004368811_003

Gualmini, A., & Moscati, V. (2009). The early steps of modal and negation interactions: Evidence from child Italian. In E. O. Aboh, E. van der Linden, J. Quer, & P. Sleeman (Eds.), *Romance languages and linguistic theory* (pp. 131–144). John Benjamins. https://doi.org/10.1075/rllt.1.07gua

Gualmini, A., & Schwarz, B. (2009). Solving learnability problems in the acquisition of semantics. *Journal of Semantics*, *26*(2), 185–215. https://doi.org/10.1093/jos/ffp002

Hickmann, M., & Bassano, D. (2016). Modality and mood in first language acquisition. In J. Nuyts & J. van der Auwera (Eds.), *The Oxford handbook of modality and mood* (pp. 430–447). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199591435.013.20

Hirst, W., & Weil, J. (1982). Acquisition of epistemic and deontic meaning of modals. *Journal of Child Language*, *9*(3), 659–666. https://doi.org/10.1017/S0305000900004967

Iatridou, S., & Zeijlstra, H. (2013). Negation, polarity, and deontic modals. *Linguistic Inquiry*, *44*(4), 529–568. https://doi.org/10.1162/LING_a_00138

Jasbi, M., McDermott-Hinman, A., Davidson, K., & Carey, S. (2021). Parents' and children's production of English negation. In D. Dionne, & L.-A. V. Covas (Eds.), *Proceedings of the 45th Annual Boston University Conference on Language Development* (pp. 360–373). Cascadilla Press.

Koring, L., Meroni, L., & Moscati, V. (2018). Strong and weak readings in the domain of worlds: A negative polar modal and children's scope assignment. *Journal of Psycholinguistic Research*, *47*(6), 1193–1217. https://doi.org/10.1007/s10936-018-9573-8

Leahy, B., & Zalneiriunas, N. (2021). Might and might not: Children's conceptual development and the acquisition of modal verbs. In N. Dreier, C. Kwon, T. Darnell, & J. Starr (Eds.), *Proceedings of the 31st Semantics and Linguistic Theory Conference* (pp. 426–445). Brown University.

Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences*, *24*(1), 65–78. https://doi.org/10.1016/j.tics.2019.11.004

Lidz, J., & Musolino, J. (2002). Children's command of quantification. *Cognition*, *84*(2), 113–154. https://doi.org/10.1016/S0010-0277(02)00013-6

Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, *154*, 40–48. https://doi.org/10.1016/j.cognition.2016.05.012

Moscati, V., & Crain, S. (2014). When negation and epistemic modality combine: The role of information strength in child language. *Language Learning and Development*, *10*(4), 345–380. https://doi.org/10.1080/15475441.2014.880640

Moscati, V., & Gualmini, A. (2007, May 11–13). More facts that isomorphism cannot explain. In T. Friedman & M. Gibson (Eds.), *Proceedings of the 17th Semantics and Linguistic Theory Conference (SALT 17), University of Connecticut* (pp. 202–209). CLC Publications.

Moscati, V., Romoli, J., Demarie, T. F., & Crain, S. (2016). Born in the USA: A comparison of modals and nominal quantifiers in child language. *Natural Language Semantics*, *24*(1), 79–115. https://doi.org/10.1007/s11050-015-9120-1

Moscati, V., Zhan, L., & Zhou, P. (2017). Children's on-line processing of epistemic modals. *Journal of Child Language*, *44*(5), 1025–1040. https://doi.org/10.1017/S0305000916000313

Musolino, J. (2006). On the semantics of the subset principle. *Language Learning and Development*, *2*(3), 195–218. https://doi.org/10.1207/s15473341lld0203_3

Musolino, J., Laity d'Agostino, K., & Piantadosi, S. (2019). Why we should abandon the semantic subset principle. *Language Learning and Development*, *15*(1), 32–46. https://doi.org/10.1080/15475441.2018.1499517

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development*, *11*(3), 191–214. https://doi.org/10.1080/15475441.2014.905169

Papafragou, A. (1998). The acquisition of modality: Implications for theories of semantic representation. *Mind & Language*, *13*(3), 370–399. https://doi.org/10.1111/1468-0017.00082

Pea, R. D. (1980). The development of negation in early child language. In D. R. Olson (Ed.), *The social foundations of language and thought: Essays in honor of Jerome S. Bruner* (pp. 156–186). W. W. Norton.

Phillips, J., & Kratzer, A. (2024). Decomposing modal thought. *Psychological Review*, *131*(4), 966–992. https://doi.org/10.1037/rev0000481

Piantadosi, S. (2011). *Learning and the language of thought*. Massachusetts Institute of Technology.

Pinker, S. (1984). *Language learnability and language development*. Harvard University Press.

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rasin, E., & Aravind, A. (2021). The nature of the semantic stimulus: The acquisition of every as a case study. *Natural Language Semantics*, *29*(2), 339–375. https://doi.org/10.1007/s11050-020-09168-6

Redshaw, J., & Suddendorf, T. (2016). Children's and apes' preparatory responses to two mutually exclusive possibilities. *Current Biology*, *26*(13), 1758–1762. https://doi.org/10.1016/j.cub.2016.04.062

Reuter, T., Feiman, R., & Snedeker, J. (2018). Getting to no: Pragmatic and semantic factors in two- and three-year-olds' understanding of negation. *Child Development*, *89*(4), e364–e381. https://doi.org/10.1111/cdev.12858

Robinson, E. J., Rowley, M. G., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006). Children's sensitivity to their own relative ignorance: Handling of possibilities under epistemic and physical uncertainty. *Child Development*, *77*(6), 1642–1655. https://doi.org/10.1111/j.1467-8624.2006.00964.x

Robinson, E. J., & Whittaker, S. J. (1986). Children's conceptions of meaning-message relationships. *Cognition*, *22*(1), 41–60. https://doi.org/10.1016/0010-0277(86)90011-9

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, *153*, 6–18. https://doi.org/10.1016/j.cognition.2016.04.006

Somerville, S. C., Hadkinson, B. A., & Greenberg, C. (1979). Two levels of inferential behavior in young children. *Child Development*, *50*(1), 119–131. https://doi.org/10.2307/1129048

Staniszewski, F., Stacey, R., & Aravind, A. (2023). Overly strong interpretations of modal verbs in child language. *Proceedings from the 47th Annual Boston University Conference on Language Development (BUCLD 47)* (pp. 715–728). Cascadilla Press.

Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development*, *59*(3), 703–718. https://doi.org/10.2307/1130570

Téglás, E., & Bonatti, L. L. (2016). Infants anticipate probabilistic but not deterministic outcomes. *Cognition*, *157*, 227–236. https://doi.org/10.1016/j.cognition.2016.09.003

Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(48), 19156–19159. https://doi.org/10.1073/pnas.0700271104

Turan-Küçük, E. N., & Kibbe, M. M. (2024). Three-year-olds' ability to plan for mutually exclusive future possibilities is limited primarily by their representations of possible plans, not possible events. *Cognition*, *244*, Article 105712. https://doi.org/10.1016/j.cognition.2023.105712

Ünal, E., & Papafragou, A. (2016). Production–comprehension asymmetries and the acquisition of evidential morphology. *Journal of Memory and Language*, *89*, 179–199. https://doi.org/10.1016/j.jml.2015.12.001

van Dooren, A., Dieuleveut, A., Cournane, A., & Hacquard, V. (2022). Figuring out root and epistemic uses of modals: The role of the input. *Journal of Semantics*, *39*(4), 581–616. https://doi.org/10.1093/jos/ffac010

van Dooren, A., Tulling, M., Cournane, A., & Hacquard, V. (2019). Lexical aspect and modal flavor in Dutch. *Proceedings from the 43rd Annual Boston University Conference on Language Development (BUCLD 43)* (pp. 203–216). Cascadilla Press.

Veselinović, D., & Cournane, A. (2020). The grammatical source of missing epistemic meanings from modal verbs in child BCS. In T. Ionin & J. E. MacDonald (Eds.), *Formal approaches to Slavic linguistics* (Vol. 26, pp. 417–436). Michigan Slavic Publications.

Von Fintel, K., & Gillies, A. S. (2010). Must… stay… strong! *Natural Language Semantics*, *18*(4), 351–383. https://doi.org/10.1007/s11050-010-9058-2

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245

Received October 10, 2024
Revision received June 4, 2025
Accepted July 7, 2025 ■