Mapping modal verbs to meanings:
an elicited production study on 'force' and 'flavor' with young preschoolers

Word Count (Abstract): 150
Word Count (Main, excl. of references): 10, 987

Abstract:

Modals (e.g., *can*, *must*) vary along two dimensions of meaning: "force" (i.e., possibility or necessity), and "flavor" (i.e., possibilities relative to knowledge (epistemic), goals (teleological), or rules (deontic)...). Comprehension studies show that children struggle with both force and flavor dimensions of modals. However, given the complex one-to-many mappings from forms to meanings, it is not clear what force or flavor children assign to the modals being tested. In this study, we use a sentence-repair task to test which modals 3- and 4-year-old children themselves prefer to produce in teleological (goal-oriented) and epistemic (knowledge-based) possibility and necessity contexts, and how these preferences differ from those of adults. Our results give a first controlled look at which modals children use to express the major flavor and force dimensions of modal verb meanings. We shed new light on children's modal systems, and show that learners generally distinguish modal flavors, but struggle distinguishing forces.

1. Introduction, Background and Hypotheses

Modal reasoning – thinking about possibilities relative to various backgrounds (e.g., knowledge, rules, desires) – is easily among the most complex types of thinking humans do, and has therefore been of central interest in developmental cognitive sciences (e.g., Byrnes & Duff, 1989; Cummins, 1996; Dack & Astington, 2011; Moscati et al., 2017; Noveck, 2001; Ozturk & Papafragou, 2015; Shtulman & Phillips, 2018). Most studies examining modal reasoning in preschoolers rely on modal language, mainly using English modal verbs (e.g., *can, must, may, have to*). Children's non-adult behaviors in these modal reasoning studies are regularly attributed to either conceptual or pragmatic deficits. However, modals are quintessential "hard words" to learn (Gleitman et al., 2005), not only because they express unobservable, highly abstract meanings, but also because of their complex grammatical patterns both within and across languages (Kratzer, 2012; Palmer, 2001; Portner, 2009). Modal word learning poses its own unique challenges at the foundational grammatical level of mapping forms to meanings. Thus, while children spontaneously produce the full range of modal verbs of English by about age 3 (Cournane, 2021; Hickmann & Bassano, 2016; Shatz & Wilcox, 1991; van Dooren et al., 2017, *to appear*), from a word learning perspective it is unclear whether it is safe to assume that 3 or even 4-year-olds have adult-like semantics for the modal words in their lexicons. In this study, we hone in on modal word learning in a novel way, using elicited production to explore how preschool English-learning children use the modals in their lexicons to express controlled modal meanings, and how their usage choices and patterns compare to those of adults from the same speech community.

Modals (e.g., *can*, *must*) vary along two dimensions: "force" (i.e., whether they express possibility or necessity), and "flavor" (i.e., whether they express possibilities relative to a body of knowledge (epistemic), a goal (teleological), or rules (deontic), among others). Possibility situations leave more than one option open, while necessity situations leave only one. Previous developmental studies show that children struggle with both the force and flavor dimensions of modal meaning (Cournane & Pérez-Leroux, 2020; Noveck, 2001; Ozturk & Papafragou, 2015; Shtulman & Phillips, 2018). Are children's difficulties due to conceptual or pragmatic difficulties, as is standardly argued, or might they stem from not knowing the force or flavor grammatical mappings of the modal words tested in these studies?

Our aim is to help tease apart conceptual and grammatical explanations for children's behaviors with modal force and flavor, and especially to gain more primary linguistic data (i.e., in our study participants are providing modals of their choosing, not responding to modals of our choosing) to militate on the issue of how young children use their modals, and how that can inform us on what conceptual cuts they make (or not) in their modal meaning space. Natural production data is heavily skewed to root uses, even more so for small children than adults (van Dooren et al., 2017, *to appear*). However, this lack in natural production could be due to a lack of propensity or opportunity to use epistemic modals, rather than missing them altogether in their lexical (or conceptual) inventory. Moreover, naturalistic production studies are further limited in that it isn't always easy to tell what meanings a child intends to convey in any given situation. For these reasons, we use a novel elicited-production method (building on Cournane, 2014) that allows us to learn more about young children's modal systems, and their usage preferences, in a controlled experimental setting with clearly defined modal situations. We specifically ask: (a) Do children distinguish force (possibility, necessity) and flavor (root, epistemic) in their modal usages by situation? If so, that suggests they have mapped those modals to differentiated concepts in form-meaning mapping. And, (b) How do children's modal preferences by modal situation differ from those of adult controls? Patterns of difference can help us better understand how children use

modals contrastively (e.g., different modals for possibility than necessity) or not (e.g., same modals for possibility and necessity), and whether that is consistent with adult-like usage, or suggestive of developmental non-adult stages researchers should be aware of if using modals as stimuli or interpreting prior studies that did so.

Let's illustrate the learning challenges for modal form-meaning mapping using a selection of English modal verbs.[1] The English modal system consists of a large number of different modals, with the commonly used set of "modal verbs" forming a grammatically distinct set (e.g., Brennan, 1993; Ramchand, 2018; Roberts, 1985; Ross, 1967). This modal system is particularly challenging from an acquisition standpoint, because the same modal meaning can be expressed by different modal words, and the same modal word can express different modal meanings. First, several modal verbs express the same "force" (1): *can, could, might,* and *may* all express POSSIBILITY (1a); the sentence in (1) doesn't rule out that Chiara is somewhere other than the office. *Should, must,* and *have to* all express NECESSITY (1b); the sentence in (2) rules out other possible locations. Moreover, many of these can express different "flavors" of possibility or necessity (2): for example, *may* can express a deontic permission (2a), or an epistemic (knowledge or evidence-based) possibility (2b); *must* can express an obligation (3a) or an epistemic necessity (3b). Flavor is standardly split into two major categories, on grammatical and conceptual bases (see Hacquard, 2011; Kratzer, 2012; Portner, 2009): ROOT (all flavors linked to possibilities given particular circumstances and preferences; e.g., ability, teleological, deontic) and EPISTEMIC (those linked to possibilities arising from knowledge or evidence). Finally, the same force-flavor combinations (e.g., epistemic possibility) can often be expressed by multiple forms (4).

(1)   a. Chiara **can/could/might/may** be in the office.   POSSIBILITY
                'It is possible that Chiara is in the office.'
        b. Chiara **should/must/has to** be in the office.   NECESSITY
                'It is necessary that Chiara is in the office.'

(2)   Sarah **may** have a plus one at this wedding…
        a. … *the invitation permitted her to bring a date.*   ROOT (Deontic)
        b. … *I saw her with someone I didn't recognize.*   EPISTEMIC

(3)   Ioana **must** eat meat…
        a. … *her doctor said so.*   ROOT (Deontic)
        b. … *I remember her ordering a steak.*   EPISTEMIC

(4)   [Doorbell rings] That **could/might/may** be Maxime at the door.
                                                        EPISTEMIC POSSIBILITY

Modals give rise to another meaning overlap mapping problem, namely a "subset problem" (Berwick 1985; Gualmini & Schwarz, 2009; Piantadosi, 2011; Rasin & Aravind, 2020; Xu & Tenenbaum, 2007, Dieuleveut et al, to appear): possibility modals are logically true in necessity situations (e.g., epistemic necessity entails epistemic possibility: if something *must* be true, it necessarily also *could* be true). This creates a further learning challenge, as children need to somehow figure out which of their modals express necessity, and which express mere possibility.

---

[1] We use the term "modal verb" to cover both modal auxiliaries like *can* or *must* and semi-auxiliaries like *have to*, which resemble verbs syntactically, but behave like modal auxiliaries semantically.

Note that possibility statements may pragmatically imply that a stronger necessity statement does not hold: otherwise, the speaker would have used it instead. This scalar implicature arises because English modals distinguish force and come in force 'duals' like <*could, must*>) (Horn, 1972) (5).

(5)   You **can** swim nude at this beach.   <CAN, HAVE TO>
          ⤑ *You don't **have to**, it's optional*   MODAL SCALAR IMPLICATURE
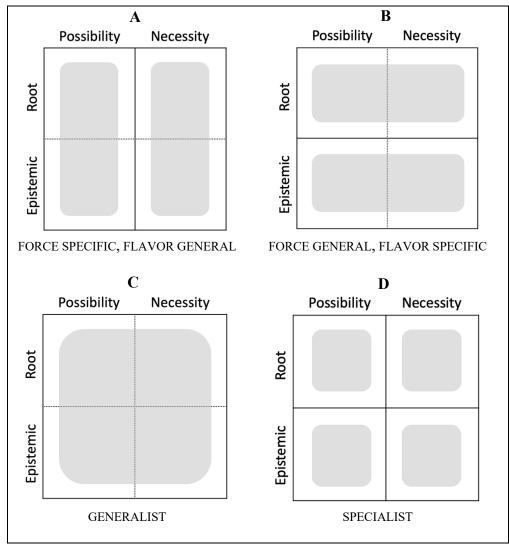
From this illustration we see that modal verbs exist in a complex system, with meanings overlapping – not to mention overlap with other grammatical categories: attitude verbs like *want*, *think* or *know* (de Villiers, 2007; Hacquard & Lidz, 2016; Shatz et al., 1983), or modal adverbs like *maybe* or *probably* (Cournane, 2021; Lassiter, 2010).

　　　In prior studies on modal force, children tend to accept possibility modals in contexts where adults prefer necessity modals (e.g., Noveck 2001; Ozturk & Papafragou, 2015). Children's over-acceptance of possibility modals in necessity contexts is often blamed on pragmatic difficulty computing the relevant implicature (i.e., understanding that use of *can* implies *doesn't have to*) (Noveck, 2001, i.a.). Adults also accept possibility modals in necessity contexts under task circumstances which encourage logical interpretations of the modals (see Ozturk & Papafragou, 2015: Experiment 1). More surprisingly from an adult standpoint, children also tend to accept necessity modals in possibility contexts, something that adults almost never do (Ozturk & Papafragou, 2015; Moscati et al., 2017). This behavior is often blamed on conceptual difficulty reasoning under epistemic uncertainty (Ozturk & Papafragou, 2015; Moscati et al., 2017; see also Leahy & Carey, 2020), and linked to a broader literature showing that young children struggle with maintaining two (or more) open possibilities (Green, 1979; Piéraut-Le Bonniec, 1980, i.a.). In Ozturk & Papafragou (2015: Experiment 1) children accept necessity modals about half the time in possibility contexts, as when a character is hiding in one of two closed boxes. They argue that children recognize that there are two open possibilities (see e.g., Fabricius et al., 1987), but perform "Premature Closure" (Acredolo & Horobin, 1987), that is, they randomly eliminate one of the possibilities. This random elimination of one of the boxes would make the *have to* statement (e.g., *The cow has to be in the green box*) true half the time: if the participant randomly eliminated the mentioned box, it would be false, but if they eliminated the unmentioned box it would be true. However, children's over-acceptance of necessity modals in possibility contexts could also be due to their lack of knowledge of the modals' underlying force, or even their interpretation of *have to* as a root necessity (e.g., an obligation to be in the mentioned box, rather than a likelihood). Do children know which modals in their input express which forces and flavors?

　　　Children's understanding of modal flavor has mostly been probed via naturalistic production corpus studies. However, these kinds of studies are limited both by what children happen to talk about, and further because it isn't always easy to tell what meanings a child intends to convey in a given situation. Putting this important caveat aside, prior studies suggest that English-learning children tend to spontaneously produce modal verbs with (what annotators judge to be) epistemic flavor up to a year later than non-epistemic flavors (often subsumed under the label "root", Hoffmann, 1966) (Cournane, 2021; Hickmann & Bassano, 2016; Papafragou, 1998; Stephany, 1979; van Dooren et al., 2017; Wells, 1979, i.a.), and perhaps even later in other languages (French: Cournane & Tailleur, 2020; BCS: Veselinović & Cournane, 2020; Dutch: van Dooren et al., 2019). This asymmetry could have various sources: conceptual (e.g., Papafragou, 1998; Stephany, 1979; Sweetser, 1990), grammatical (e.g., Cournane, 2021; Veselinović & Cournane, 2020), or differences in input frequency (e.g., van Dooren et al., 2017). Do children

know which modals in their input can express which flavors? As children's modal input is highly skewed toward root, they could be slow in realizing that modals like *could, must* or *have to* can express epistemic modality. Alternatively, children could have difficulty differentiating "flavors" of modality conceptually, and assume initially a more general concept of possibility or necessity (cf. Shtulman & Phillips, 2018).

Throughout this paper we use 2 x 2 maps to illustrate the "modal meaning space" (based on Vander Klok, 2012), as in Figure 1, and to plot modal usage patterns by usage situations. The meaning space (and associated usage situations) are defined by the two major forces (x-axis; possibility and necessity) and two major flavors (y-axis; root and epistemic). Figure 1 shows 4 patterns for how modal elements may express the modal meaning space: modal systems may show usage patterns, relative to situations, such that: (A) force is lexically distinguished, but flavor is not, (B) flavor is lexically distinguished, but force is not, (C) neither force nor flavor are lexically distinguished, or (D) both force and flavor are lexically distinguished.

Figure 1. Four possible ways that modal words may map to the 2 x 2 modal meaning space.
Map A: Distinguish Forces: modal forms show consistent forces, but are flavor-general.
Map B: Distinguish Flavors: modal forms show consistent flavors, but are force-general.

Map C: General Modals: modal forms are generalized to all major modal meaning combinations.
Map D: Distinguish Force & Flavor: modal forms are specialized to specific force-flavor pairings.

Languages differ in how their modal forms lexicalize this modal meaning space (Nauze, 2008; Vander Klok, 2012). For example, English modal verbs distinguish force, but they can be used to express different flavors (1)-(4) – i.e., *must* is always necessity (contrasting with e.g., *could*), but it can express either root or epistemic flavors, making English modal verb use best captured by Map A. Typological work, based on careful semantic fieldwork (Bochnak & Matthewson, 2015; Matthewson, 2004), has shown that the other three maps exist as well: some languages have modals used for both force situations, but always in the same flavor (Map B; e.g., Nez Perce, Deal, 2011), others have a sort of all-purpose modal element used for both major forces and major flavors (Map C; e.g., Washo, Bochnak, 2015)[2], and others rely on specific lexical items for each major cell (Map D; e.g., Paciran Javanese, Vander Klok, 2012)[3]. Large corpora of English child directed speech (van Dooren et al., *to appear*) show that while the English modal verb system basically aligns with Map A, there are specific modal verbs that function more like Map D – for example, unnegated *can* is almost exclusively used in root possibility contexts, and *have to* in root necessity contexts. Likewise, *might* and *must* is mostly used in epistemic contexts (root necessity *must* seems formal to most speakers, see e.g., Tagliamonte & D'Arcy, 2007). Here we can take advantage of English's relatively rich and diverse modal verb vocabulary to probe if children distinguish forces and flavors by using different words.

In this study, we use a child-friendly sentence-repair task (Cournane, 2014) to investigate which modal verbs children at the onset of a productive modal system (3- and 4-year-olds) prefer to use for each of the meaning cells of the 2x2 Modal Meaning Space.[4] We use short stories illustrating each force-flavor combination in a maximally comparable way, and provide children the opportunity to use modals of their own choosing to describe each meaning cell. For epistemic stories we use a hiding game with two boxes as possible hiding spots, with both boxes closed for possibility conditions and one box open and shown to be empty for necessity conditions (as in Ozturk & Papafragou, 2015). For root stories, we use teleological (goal-oriented) root modality, with two possible roads leading to the goal destination (see also Dieuleveut et al., 2021). We chose teleological modality as representative of root modalities because it is readily imageable, and does not require the added cognitive (and task) complications of a body of rules or authority figure, as with deontic root modality (see Dack & Astington, 2011; Hirst & Weil, 1982). In possibility conditions both roads are open, while in necessity conditions one road is blocked by construction. We also test adults from the same community. Do children distinguish flavor by using different words for root and epistemic modalities? Do they distinguish force by using different words for

---

[2] Maps B & C - This is distinct from how English-type possibility modals (e.g., *can*, *might*) may be used in necessity situations under certain circumstances, as they are logically true (necessity entails possibility). English modals come in "duals", such that use of *can* can imply that a stronger (=necessity) modal is not true (Ex5). Importantly, not all languages have such modal 'duals'. The Sahaptian language Nez Perce, spoken in the Northwestern US, for instance, only has a possibility modal, which speakers use to describe both possibility and necessity situations (Deal, 2011). St'át'imcets (Lillooet Salish), spoken in British Columbia, only has necessity modals (Rullmann et al., 2008). Languages with modal systems that show Map B patterns lack force duals (at least within flavors).

[3] Notice that in English too there are words or expressions with modal meanings, but not from the "modal verb" set, that can differentiate each cell, in line with Map D; crucially, Paciran Javanese grammatically only makes use of fully force-flavor dedicated modals.

[4] Note that if a child only produces one modal verb in a given condition (e.g., *can* in all four teleological possibility trials), it doesn't necessarily follow that they haven't learned that *could*, for example, can be used to express teleological possibility meanings as well. But, it does indicate which modal word they prefer in that context.

possibility and necessity? How do children use modals words in contrast (or not) with each other? How like adults' patterns of usage are children's patterns of usage?

Our hypotheses track the maps laid out in Figure 1, inspired by typologies discovered through cross-linguistic fieldwork, and stated in terms of usage patterns for modal forms to modal situations. Hypotheses differ only along the dimension of whether (or not) the modals used as a set distinguish modal force and/or flavor:

- HYPOTHESIS A: MODALS USED DISTINGUISH FORCE, BUT NOT FLAVOR (MAP A)
- HYPOTHESIS B: MODALS USED DISTINGUISH FLAVOR, BUT NOT FORCE (MAP B)
- HYPOTHESIS C: MODALS USED DISTINGUISH NEITHER FORCE NOR FLAVOR (MAP C)
- HYPOTHESIS D: MODALS USED DISTINGUISH BOTH FLAVOR AND FORCE (MAP D)

Our methods directly probe which modal words children prefer for different forces and flavors, but they also allow us to probe indirectly, which conceptual cuts they make: if children use different words for different forces or flavors (e.g., for force for Map A) this suggests that the conceptual distinction is likely also made. For example, if children use different modals for possibility versus necessity, it suggests they have mapped those modal meanings separately. The use of possibility modals in both possibility and necessity scenarios is truth conditionally correct, even if not always optimal from a pragmatic perspective. For this reason, observing a force contrast will critically involve observing necessity modals used exclusively for necessity scenarios. Importantly, observing a lack of distinction does not allow us to conclude anything about its source– for example, if children do not distinguish possibility from necessity, we cannot conclude over whether that is for conceptual reasons (in line with some theories in developmental psychology, e.g., Acredolo & Horobin, 1987) or linguistic-mapping reasons (as is true for conceptually mature adult speakers of languages that don't lexically contrast them).

Given description and corpus work on adult English (Brennan, 1993; Kratzer, 2012; Ramchand, 2018; van Dooren et al., 2017, among many others), we expect our English-speaking adult participants to use modal verbs in line with Hypothesis A (English modal verbs are flavor-variable, but fixed for force) with perhaps some usage patterns suggestive of Map D (some English modal verbs also show flavor-specific preferences in spontaneous speech corpora, e.g., *can* as root possibility, *might* as epistemic possibility (van Dooren et al., 2017).

For children, this study is more exploratory – we know what modals they produce in spontaneous corpora from approximately 2-4 years old, but whether they have the adult-like lexical contrasts (Map A), is the open question we set out to explore in a controlled manner. Prior work in developmental psychology, about the development of possibility reasoning, suggests preschoolers are not fully adult-like in this area (for overview, Leahy & Carey, 2020), and likewise, in 2-3-year-olds modal verb usage by force appears adult-like for possibility modals but not necessity modals (Dieuleveut et al., to appear). For flavor, developmental psychology work is encouraging that toddlers and preschoolers can reason about their own and others' knowledge states (e.g., Gergely et al., 1995; Goupil et al., 2016; Onishi & Baillargeon, 2005; Southgate et al., 2007), as required for epistemic reasoning. And, it is also encouraging for various root modalities (e.g., Dunn, 1988; Tomasello & Carpenter, 2007; Woodward, 1998). Likewise, modal usage appears to show an array of modal flavors, especially from age 3 (Cournane, 2021; Papafragou, 1998), even including some child samples showing the same modal (e.g., *must*) used in seemingly distinct root and epistemic situations from age 2 or 3 (van Dooren et al., to appear).

With this in mind, we expect child modal usage patterns to support Hypothesis B, patterning with Map B. Map B would suggest a lack of distinction by force, but a distinction by flavor. Such a distinction by flavor could be due to children not realizing that English modal verbs are flavor-variable, and thus treat modal verbs as only expressing one flavor.

Broadly, our results show that our expectations were borne out. Adults behave as expected for English speakers (Map A, aspects of Map D), primarily supporting Hypothesis A. Three and four-year-old children's modal usage patterns reveal that they tend to be force generalists (not showing contrastive usage for force), but flavor specialists (showing contrastive usage for flavor), best supporting Hypothesis B. More specifically, for force: children use possibility modals in both possibility and necessity contexts, like adults do sometimes (they are logically true in both, so this is unsurprising), but children also use necessity modals in possibility contexts, which adults almost never do. For flavor, children overwhelmingly use different modals for different flavors: for example, they tend to prefer *can* for teleological modality, and *might* for epistemic modality (showing flavor-specific patterns), unlike adults who tend to use modals across flavors, especially *could* in both teleological and epistemic possibility contexts.

Children's non-adult behaviors with force could have conceptual (Ozturk & Papafragou, 2015) or, as we also suggest here, also linguistic bases. For example, children may be struggling to figure out which of the modals in their language are necessity modals given the subset problem (Dieuleveut et al., to appear), namely the fact that necessity entails possibility. We can see that children do distinguish modal flavor lexically: this suggests that they discriminate the two major flavors of possibilities conceptually (root, epistemic), and know appropriate modals to describe them (this finding is consistent with corpus work). However, they may not yet appreciate fully that English modal verbs are flavor-variable, since both as a group, and when we look at individual responses, they tend to use specific modals in only one flavor; this is a grammatical issue, perhaps related to biases for one-to-one form-meaning mapping, which may be overcome through bootstrapping from tense and aspectual information (which differentiates root and epistemic modality, albeit indirectly; van Dooren et al. to appear).

2. Methods

To see which modals children prefer for different modal situations, we adapted Cournane's (2014) modal sentence repair task to elicit modal verb productions. Cournane played participants short stories with test sentences spoken by Elmo, but where the modal verb was obscured by Elmo's dog barking (e.g., "She <woof> wear ballet slippers"). Participants were prompted to repair the sentence with "Oh no, Zappy barked – what did Elmo say?". She measured participants' responses in 5 conditions based on modal flavor (ability, deontic, teleological, future, epistemic). She did not control for force of the contexts (possibility vs. necessity), as her main goal was syntactic: to observe category preferences of modal verbs (semi-auxiliary vs. auxiliary) supplied by children (age 5) and adults. To address our research questions, we have adapted her sentence repair task to test children as young as 3-years-old on which modals they prefer to use in four types of modal situation, explicitly crossing force and flavor: teleological possibility, teleological necessity, epistemic possibility and epistemic necessity contexts (as in Figure 1).

2.1. Stimuli

We created short, illustrated stories to establish the modal situations for our critical sentences. Audio for each story was pre-recorded using Praat (Boersma, 2001), and accompanying illustrations were created using Adobe Draw. Stimuli were presented using Microsoft PowerPoint,

with animations added as needed for clarity or emphasis. Participants were trained to repeat pre-recorded sentences to a shy snail puppet, Mr. Drooly, who doesn't always hear what is happening because he hides in his shell. In critical sentences, pink noise was spliced into the sentence where it would be natural for an adult to use a modal verb. Each critical sentence was recorded with the nonsense word *gorp* in place of an actual modal verb. After recording, the *gorp* was manually removed using Praat and replaced with 0.45 seconds of pink noise. This kept any potential co-articulation cues consistent and uninformative in each critical sentence. All materials are available at [REDACTED].

Teleological and Epistemic stories are designed to be minimally different, so as to allow direct comparison between the two flavor conditions. Before starting the teleological trials, the experimenter establishes that a character, Kat, has the broader goal of retrieving items from different stores—an explicitly goal-oriented context. In each story, Kat's goal location appears opposite her with two distinct paths in between, representing the two means of achieving her goal.

Figure 2 provides sample teleological possibility (left column) and teleological necessity (right column) stories. These stories exemplify the top row of our 2 x 2 modal meaning space maps (Figure 1). The trial type determines whether or not both paths remain accessible to Kat: in the possibility condition, both paths remain open to her throughout the trial. In the necessity condition, a construction gate and traffic cones appear, blocking one path, and leaving only one possible route open for Kat to reach her goal. The color of each path and the location of the blocked path (top or bottom) vary systematically throughout the experiment. The frame of the critical sentence supports a teleological modal by using an eventive predicate (*go down the path*)[5]. The syntactic frame is identical in both teleological possibility and necessity conditions, but two recordings were made, one for possibility and one for necessity, to keep the prosody as natural and supportive as possible.

---

[5]Root and epistemic modals tend to combine with different predicates: eventives for the former, and statives for the latter (Condorvadi, 2002; van Dooren et al., to appear; Ramchand, 2018). To further encourage the right modal interpretation, we used eventive predicates for teleological conditions, and stative predicates for the epistemic ones.

Teleological possibility		Teleological necessity	
	Now Kat is going to the library to get books!		Now Kat is going to the bakery to get a cake!
	There are two ways to get to the library. One way is to go down the yellow path...		There are two ways to get to the bakery. One way is to go down the red path...
	Look! Or...		But uh-oh – it's blocked! So...
	Kat <<pink noise>> go down the brown path!		Kat <<pink noise>> go down the pink path!

Figure 2. Sample teleological (goal-oriented root modality) trials, illustrating the root possibility (left) and root necessity (right) conditions.

Figure 3 provides sample epistemic possibility (left column) and teleological necessity (right column) stories, which we modeled after Experiment 1 in Ozturk & Papafragou (2015). These stories exemplify the bottom row of our 2 x 2 modal meaning space maps (Figure 1). The experimenter introduces the epistemic trials by explaining that Nick is playing hide-and-seek in different boxes. Each trial starts with two empty distinctly-colored boxes, which represent the two potential hiding spots for Nick. A curtain lowers while Nick hides. When the curtain lifts, Nick has vanished, and both boxes are closed. In possibility trials, both boxes stay closed, so there are two possible locations for Nick's true whereabouts. In necessity trials, one box pops open to reveal that it's empty. In this case, it's necessary that Nick is hiding in the other box. Nick's true location is never revealed because we wanted children to think that the game was about describing where Nick may be, rather than about finding Nick, as that would make using a modal verb to describe the scene less appropriate. The color and location (left or right) of the box which pops open was counterbalanced. The frame of the critical sentence supports an epistemic modal by using a *be* progressive (*be hiding in the box*), and is identical in possibility and necessity trials. Two recordings were made, one for possibility and one for necessity, to keep the prosody as natural and supportive as possible.
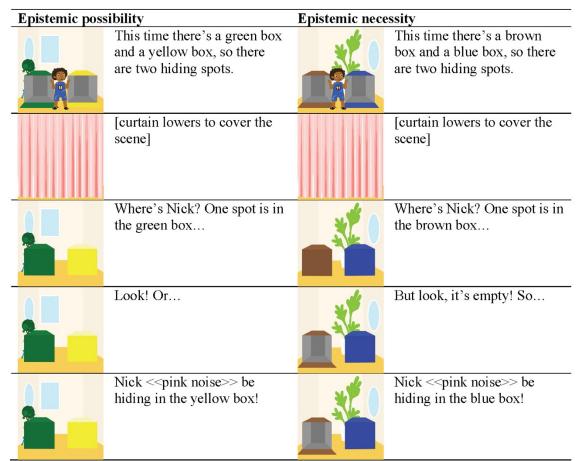
Epistemic possibility		Epistemic necessity	
	This time there's a green box and a yellow box, so there are two hiding spots.		This time there's a brown box and a blue box, so there are two hiding spots.
	[curtain lowers to cover the scene]		[curtain lowers to cover the scene]
	Where's Nick? One spot is in the green box…		Where's Nick? One spot is in the brown box…
	Look! Or…		But look, it's empty! So…
	Nick <<pink noise>> be hiding in the yellow box!		Nick <<pink noise>> be hiding in the blue box!

Figure 3. Sample epistemic trials, illustrating the epistemic possibility (left) and epistemic necessity (right) conditions.
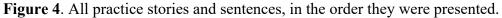
2.2. Procedure

In a quiet testing room, participants were seated next to the experimenter and told they would be listening to some stories on the computer with Mr. Drooly, a shy snail puppet operated by the experimenter. The experimenter explained that the participant's job was to listen to the story, and then retell some parts of the story to Mr. Drooly (those that happened when he was hiding in his shell from shyness). Participants were asked to listen very carefully because sometimes words were missing from the stories. Children and adults did the exact same task, except adults were told the task was designed for children. Children's responses were recorded by the experimenter using a pen and paper during the study. As a backup, video recordings were made (with parental consent) so that the pen and paper coding could be double-checked. The entire experiment took about 12 to 15 minutes.

There were 4 practice stories, including a total of 10 practice sentences (Figure 4). The first several sentences contained no "glitch" (i.e., pink noise), and then we introduced the glitch over where a quantifier (e.g., *two*, *some*) or noun (e.g., *dog*, *one*) would be. Participants were prompted to repeat each sentence to Mr. Drooly before moving on to the next one. The practice trials served four purposes, especially important for our child participants: (a) to help build confidence repeating sentences to Mr. Drooly; (b) to introduce the kinds of sentences they would hear and scenes they would see during the experiment; (c) to familiarize them with the experience that some sentences were corrupted by a "glitch" (i.e., pink noise) before they got to the critical trials; and,

(d) to provide an exclusion criterion: participants who failed to repeat and repair corrupted sentences during practice were excluded from analysis. If participants repaired corrupted sentences with a word(s) from the appropriate syntactic category, they passed.

Trial	Pre-recorded sentences	Illustrations
1	a. There is a girl and a boy! b. There is a red apple! c. There is a yellow chair! d. There is a blue ball, too!	
2	a. Look! There are two toys to play with. b. One toy is in the sandbox. Look, it's a car! And the other toy is a beach ball!	
3	a. Look! He has a black cat, and a white cat! So there are two cats. b. Where are the cats? One cat is on the rug – look! And the other cat is on the sofa.	
4	a. This time, there is a yellow dog and a red dog, so there are <<pink noise>> dogs. *Target responses: there are {two/some/the} dogs.* b. Where are the dogs? One dog is on the chair – look! The other <<pink noise>> is under the table! *Target responses: the other {dog, one} is under the table*	

Figure 4. All practice stories and sentences, in the order they were presented.

At the point when participants were supposed to repeat a practice or critical sentence, Mr. Drooly came out of his shell and asked, e.g., "Nick whaaat…?". After the participant repeated the sentence, he hid away again, and the experimenter started the next trial. If participants did not respond immediately, the experimenter waited for several seconds, and then asked them if they would like to hear the sentence again. In these cases, Mr. Drooly hid away again while the experimenter replayed the sentence. This was common with children for the first couple practice trials, but stopped once children got more confident about the task. Sentences were only ever repeated once, and the stories which set up the critical sentences were only repeated if the story had been interrupted (either by the participant or ambient noise). Participants saw all the practice trials before moving on to the critical trials.

We piloted several versions of this training, with different methods involving first a dog and then Mr. Drooly and his backstory, different numbers of plain and glitch-containing practice sentences, and different types of glitches (first the bark method from Cournane (2014), which was too distracting for our younger population, and other "colors" of noise). This combination was most successful for getting the youngest children to repeat the sentences to Mr. Drooly and repair the glitch seamlessly with a word from their own grammar.

4.4. *Experimental Design*

Modal flavor and force were fully crossed (2x2 design) to make 4 test conditions (teleological possibility, teleological necessity, epistemic possibility, epistemic necessity), presented within-subjects. Each condition had 4 trials, yielding a total of 16 test trials per participant, to give us the opportunity to observe whether participants used multiple modal verbs within a single condition.

Participants first completed the practice trials. Then all test trials of the same flavor (TELEOLOGICAL, EPISTEMIC) were presented in blocks, with block order varied between subjects; about half of the participants (n=18 children, n=12 adults) completed all the teleological trials first before moving on to the epistemic trials, and the other half, vice versa (n=20 children, n=12 adults). Possibility and necessity trials were pseudo-randomized such that no two trials of the same type occurred more than twice in a row. For some participants (n=21 children, n=11 adults), the first trial in the experiment was a necessity trial, and for some (n=17 children, n=13 adults), it was a possibility trial. These manipulations yielded 4 experimental orders of trial presentation.

2.3. Pre-classification of Modal Verbs

Given how English modal verbs work, we can classify them by force prior to collection. Flavor is often variable in English, so we refrain from pre-classifying by flavor, and will instead appeal to corpus research on the input to children to give softer classifications for flavor usage tendencies in our Discussion. Modal verb classification by force is given in (6). We have pruned the lists to include only those verbs that actually occurred in our study (e.g., *shall* and *ought* are English modal verbs, but they are very rare/archaic and were not actually used by speakers in the DC area). We treat future auxiliaries (*will*, *would*, *gonna*), which are grammatical as sentence repairs in our task, as a distinct category: their semantic status as modals, is debated (e.g., Kissine, 2008; Klecha, 2013); when treated as modals, they are associated with "metaphysical" modality (e.g., Condoravdi, 2002), and *will* and *would* often with necessity force (e.g., Copley, 2009; Kaufmann, 2005), but not always (e.g., Cariani & Santorio, 2018).

(6)   Modal verb classification by lexical force
         a. Possibility modals: *can*, *could*, *might*
         b. Necessity modals: *must, should, has/had to*
         c. Future auxiliaries: *will, would, gonna*

2.4. *Participants*

Child participants (n=38, mean age=4;1;14, range=3;0;13-5;3;25) included in the final sample were tested in the lab at [REDACTED], and in preschools in the surrounding communities. Our age range starts at 3-years-old because it is both the youngest age at which children could participate in the task, and the age when typically-developing English-learning children reliably show the full array of modal verbs, including with some epistemic uses thereof. Adult participants were undergraduate students (n=24, mean age=21 years, range=18-28) recruited from the [REDACTED] to establish an adult baseline for modal preferences in our task.
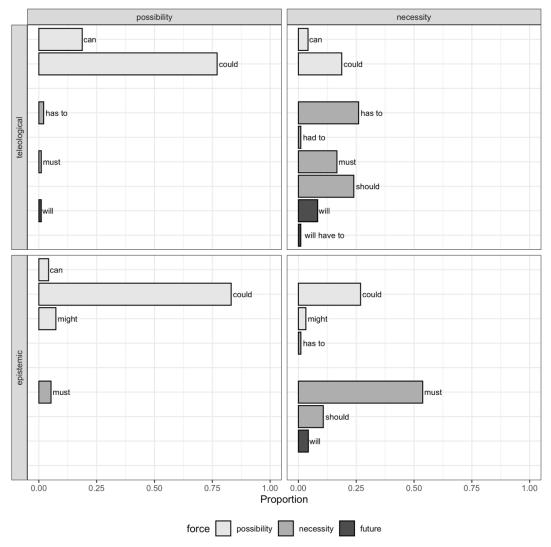
An additional 29 children were tested but excluded from analysis, for the following reasons: failing the practice trials (n=9); only responding in the present or past tense, indicating they thought the task was a guessing game, not a repeating-the-sentence game (n=8), which is a common issue for possibility-related tasks (see Ozturk & Papafragou, 2015); refusing to speak to the experimenter out of shyness (n=7, all but one aged 3), which is common for 3-year-olds; and,

for being too fidgety (n=5). This high exclusion rate (29/67 children) suggests that this kind of meta-linguistic task can be difficult for younger children to participate in, probably unsurprisingly to anyone who has worked with this age group. A Welch two-sample t-test revealed that excluded child participants (n=29, mean age=3;9;3, range=3;0;10-5;4;15) were significantly younger than those included in the study $t(60.294)=-2.6462$, $p=0.01$).

3. Results
3.1. Adult Results – Aggregate
Adult responses were 99% (380/384) modal verbs (auxiliaries and semi-auxiliaries), meaning adults consistently accurately repeated the syntactic frame that they had heard, repairing the pink noise in the expected way. Figure 5 shows adult modal verb response counts and proportions by lexeme by condition.

Figure 5. Aggregate adult modal verb response by proportion by lexeme by condition. Color indicates classification as possibility, necessity, or future.

Adults primarily use *could* (74/96, 77%)[6] and *can* (18/96, 19%) to express teleological possibility (only 4/96 exceptions). In teleological necessity, there is more variation, but adults primarily use *has to* (25/96, 26%) and *should* (23/96, 24%). In epistemic contexts, adults prefer to use *could* (79/95, 83%) to express possibility and *must* (50/93, 54%) for necessity. Summarizing, adults primarily use *could* to express both epistemic and teleological possibility, but tend to use different words to express epistemic and teleological necessity. Notably, while the possibility modal *could* appears in both possibility and necessity conditions for both flavors, necessity modals like *have to*, *should*, and *must* overwhelmingly appear only in the necessity condition. Adults very rarely use necessity modals (e.g., *have to*, *must*) in possibility situations (8/384 responses), largely ruling out Maps B and C, both of which do not distinguish by force.

With respect to our hypotheses, adult modal usage patterns by condition show that adults tend to distinguish forces, but use at least some modals in a flavor-variable way (especially clear in possibility conditions), best aligning with HYPOTHESIS A (Modals used distinguish force, but not flavor, Map A). To some extent, and especially for necessity conditions, adults use modal forms that are specialized to specific force-flavor pairings, aligning with HYPOTHESIS D (Modals used distinguish both flavor and force, Map D).

3.2. Child Results – Aggregate

Children's responses were much more variable than those of adults, both for category of response (modal verb vs. other strategies), and among modal verb lexical items. Children provided 40% (226/563) modal verb repairs for the pink noise, accurately repeating the syntactic frame they had heard in these instances.[7] We focus first on this category of child response, because it is directly comparable to the adult results. Figure 6 shows child modal verb response counts by lexeme by condition.

---

[6] Here we are giving the counts of each modal verb out of the total for each condition we mention, and the percentage that modal verb makes up for those by-condition responses.

[7] Responses with minor changes which didn't affect the frame, like saying *he* instead of *Nick*, or anaphoric *one* instead of *path* or *box*, were counted as frame-compliant.
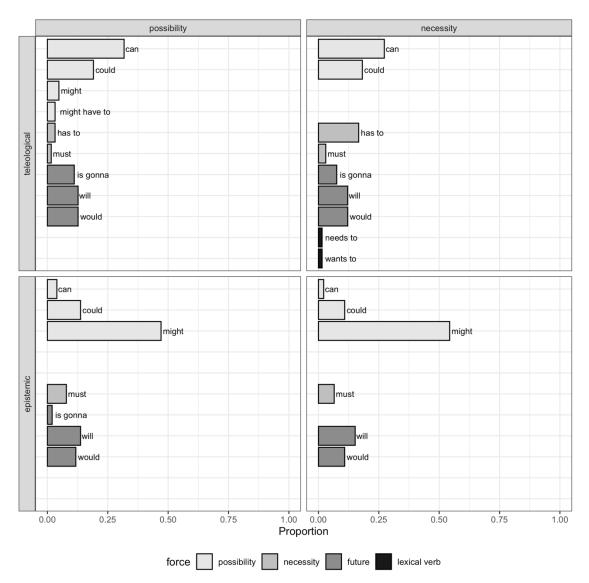
Figure 6. Aggregate child modal verb response proportions by lexeme by condition. Color indicates classification as possibility, necessity, future, or lexical verb.

Figure 6 shows that children's modal verb usage is remarkably similar across forces - they are using the same modals in similar proportions for both teleological possibility and necessity, and for both epistemic possibility and necessity. Among the modal verbs, children use *can* most often in both teleological contexts (38/129, 29%), followed by *could* (24/129, 19%), and *might* most often in both epistemic contexts (49/97, 51%). Overall, children use the same modal verbs to express possibility and necessity within each flavor, with one exception: children use *has to* more often to express teleological necessity (11/66, 17%) than teleological possibility (2/63, 3%). Children show mostly distinct modal verb usage by flavor, with the exceptions of possibility modal *could*, which occur at similar rates across all four conditions. Children also used future auxiliaries (*will*, *would*) in all four conditions, and *is gonna* primarily in teleological contexts (10 children gave at least one future response, 3 of which used them in all four conditions). Adults also gave a few instances of *will*, almost all for necessity. No adult gave *would* or *is gonna* as a response.

17

In contrast to flavor, children do not use different modal verbs to distinguish possibility and necessity in epistemic contexts. Children prefer *might* (49/97, 51%) both to express epistemic possibility (24/51, 47%) and epistemic necessity (25/46, 54%), and other modal verbs occur at similar rates across both forces. This differs from adults, who did not use *might* often, not even for epistemic possibility. In comparison, adults also used *must* frequently in epistemic necessity, but children used it infrequently and evenly distributed by force. And while adults have clear preferences for *could*, especially pronounced in both possibility conditions, children use less *could* than adults and roughly evenly distributed across all four conditions.

Considering our hypotheses, we see that children's aggregate modal verb uses show much evidence for flavor distinctions, but little to no evidence for force distinctions (with the exception of low-rate teleological uses of *has to* differing by force). The hypothesis best supported by our child results is HYPOTHESIS B (Modals used distinguish flavor, but not force (Map B), with usage of *could* also providing some support for HYPOTHESIS C (Modals used distinguish neither force nor flavor, Map C). Future auxiliary usage may also be taken to support HYPOTHESIS C , depending on one's analysis of English futures.

Other kinds of child responses (337/563, 60%) were largely systematically aligned to flavor, but not force, fitting in this general way to the pattern with modal verbs. The most frequent non-modal verb strategy occurred when participants tensed the BE auxiliary (e.g., *Nick **is** hiding in the yellow box* for *Nick ____ be hiding in the yellow box*), making the sentence an unmodalized declarative, instead of providing a modal verb before the BE auxiliary (accounting for 58% (195/337) of non-modal-verb responses). This kind of response occurred more in epistemic conditions (120/195, 64%), and was very slightly more common in necessity (105/195, 54%). The few times that adults gave non-modal-verb responses, they were also of this type in epistemic conditions (n=4/384, 3 in necessity trials, one in possibility). Probably due to this competing strategy, children gave fewer modal verb responses in the epistemic condition (97 modal verb responses) than in the teleological condition (129 modal verb responses). In contrast, the next most frequent non-modal-verb response type occurred more often in teleological trials (91/124, 73%); children sometimes simply repeated the test sentence, leaving out the pink noise and using a bare verb. For example, children said, *Kat go down the yellow path* for *Kat ____ go down the yellow path* (124/337, 27% of non-modal-verb responses; these responses were near-equally distributed by force condition; 63/124, 51% necessity).

The remaining 5% (18/337) of other responses included modal words from other syntactic categories (attitude verbs, adverbs) or modal verbs in a frame of the child's creation (Table 4). As with the more common modal verb and non-modal responses, these responses are also compatible with a teleological vs. epistemic understanding of the context. Expressions like *maybe*, *I'm sure*, *I guess…,* which express epistemic or doxastic possibility or necessity, were used almost exclusively in epistemic contexts, with one exception (*I think she goes down the <color> path*), which was used to express teleological necessity (Table 4). Expressions of desire or need were used almost exclusively in teleological possibility contexts, with one exception (e.g., *There needs to be more hiding spots*), which expresses a need that the child herself had for the hide-and-seek game to be more fun. Children occasionally changed the frame in teleological possibility contexts to talk about Kat's broader goals (e.g., *She wants to get some books!*), rather than the means of achieving them.

Possibility	Necessity
She **should** go across either one (#91)	I **think** she goes down the <color> path (#65)
Both ones **would** be fine to go down (#91)	
She **has to** go to the store (#96)	
She **wants to** get some books! (#96)	
I **guess** he's hiding in the <color> one (x3 by #83)	I **guess** he's hiding in the <color> one/box (#83)
I **guess** he's hiding in the <color> box (#96)	**Maybe** he's hiding in the <color> box (#42)
I'm **pretty sure** he's hiding in the <color> box (#80)	He's hiding in the <color> box, **maybe** (#87)
Maybe he's hiding in the <color> box (#87, 42, 49, 58)	There **needs to** be more hiding spots (#91)

Table 4. Children's modal responses that did not maintain the provided syntactic frame, by force and flavor condition. Participant numbers in parentheses. Each response type was used once unless otherwise noted.

Table 5 shows counts and proportions of possibility (e.g., *can*) versus necessity (e.g., *must*) modal verb responses for adults and children in aggregate, by each force condition and overall.[8] We use our pre-determined classifications of English modal verbs for force on the basis of semantic description (6). We treat future responses separately. Both adults and children provided mostly possibility modal responses. However, looking by condition, adults use more possibility modals in possibility conditions than necessity conditions, while children use possibility modals at roughly equal rates. Adults use proportionally many more necessity modals that children, and children also differ from adults in using many futures, at roughly equal rates for both conditions, while adults use very few futures.

		Condition Type		
	Response Force	Possibility	Necessity	Total
	Possibility	182 (0.48)	50 (0.13)	232 (0.61)
Adults	Necessity	8 (0.02)	125 (0.33)	133 (0.35)
	Future	1 (<0.01)	14 (0.04)	15 (0.04)
	Possibility	68 (0.30)	61 (0.27)	129 (0.57)
Children	Necessity	9 (0.04)	18 (0.08)	27 (0.12)
	Future	37 (0.16)	33 (0.15)	70 (0.31)

Table 5. Counts and proportions of modals by response force by force condition (Possibility, Necessity, Future). Proportions calculated relative to total number of responses for the age group: for adults (n=380) and children (n=226).

We have been careful to show futures as separate, but they are typically assumed to have universal force[9] (i.e., in line with necessity modals), therefore we ran our statistics grouping them with necessity modals. We fit two models, one on the adult data and one on the child data, to assess whether participant responses (possibility vs. necessity) differed by force condition. We expected adults would (Hypothesis A), but children would not (Hypothesis B). We used the glmer function in the statistical package lme4 in R (R Core Team 2013). The data were fitted into a generalized
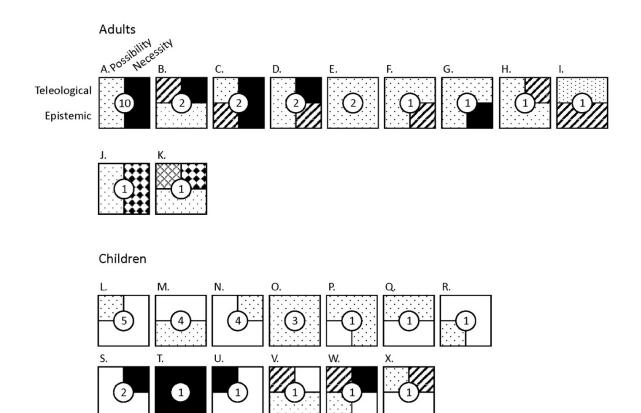
---

[8] There were a very few instances where participants used a grammatically appropriate modal cluster to repair the pink noise (*will have to, and might have to*), we classified these by the first word (future and possibility, respectively).
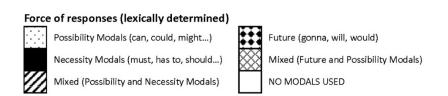[9] But see Santorio & Cariani (2018) who argue that *will* lacks quantificational force altogether.

linear-mixed (logit) model using the maximum likelihood method (Laplace Approximation) (Baayen, 2008; Dixon, 2008; Matuschek et al., 2017). The model had production of a possibility modal as the reference for the dependent variable, and possibility condition as the reference level for the force condition, and included participant as a random factor (RESPONSE_FORCE ~ FORCE_CONDITION + (1|PARTICIPANT). For the adult data, we found a significant increase in necessity responses in necessity conditions (ForceCondition_Necessity, ß = 5.20, Z = 9.70, *p* < 0.001). For the child data, we found no significant difference in the force of modal responses by the force condition (ForceCondition_Necessity, ß = 0.30, Z = 1.29, *p* = 0.20).

3.4. Individual response maps
Here we focus in on how individuals use modals across the modal meaning space, to get a sense of individual modal systems. We created 2x2 maps of the four experimental conditions for each individual, plotting their modal use by the lexical force of the modals they used per cell (Figure 7), treating futures as a distinct set. The pattern and shading in each cell correspond to the force of the auxiliary (i.e., frame-compliant) responses which were used in the relevant condition. The number of participants who share each map is indicated by the number in the center. If only possibility modals were used in a condition, we illustrate with spotted pattern, necessity with black, and a mixed set with the diagonal stripe; future use is marked with filled diamond, and mixed future and possibility with diamond lines. If no modal verbs were used in a condition, we left it white (children only).

Figure 7. Individual modal maps (Adults: A – K; Children: L – GG), plotting frame-compliant responses by their lexical force (possibility, necessity, future) for each of our 4 test conditions (labeled on Map A). Number of individuals who used a particular map is noted in the center.

Focusing on the adult maps (A-K), 14 of 24 adults used non-overlapping sets of modal verbs for possibility versus necessity contexts (Maps A and D)[10]. In the most common adult pattern (Map A; n=10), adults provided possibility modals in possibility contexts, and necessity modals in necessity contexts, across both flavors (as expected given English modal verb semantics). Map D

---

[10] Map J is in line with this pattern, except this individual used *will* in necessity contexts.

(n=2) is also consistent with this, but with some possibility modal use in epistemic necessity. Another 6 adults used necessity modals in one flavor, but not the other (Maps B; F-I). Maps B and I are inconsistent with our expectations for English speakers, and Map C is mostly consistent, but with some unexpected use of necessity modals in epistemic possibility contexts. Two adults used only possibility modals (Map E), and two adults used some future auxiliaries (Maps J & K).

Turning to the children's maps (L-GG)[11], we see that typically children used modal verbs to express one flavor or the other: 14 of 37 children used modal verbs in teleological contexts only (Maps L, N, Q, S, U, X), and 5 children used modal verbs in epistemic contexts only (Maps M, R). Nineteen children used exclusively possibility modals (Maps L-R), 4 children used exclusively necessity modals (Maps S-U), and only 3 children used a mix of both possibility and necessity modals throughout the experiment (Maps V-X) (compare to 22/24 adults who used a mix). A further 4 children used only future auxiliaries, and 6 used a mix of futures and possibility modals. Unlike most adults, children did not systematically use possibility modals in possibility conditions and necessity modals in necessity conditions; no child shows the most common adult map (A). On our inspection of individual data, we noted that most of the children who used exclusively possibility modals were under 4-years-old. Around the 4th birthday, children started using more necessity modals and futures, and some of these older children used necessity or future modals exclusively.

4. Discussion

Our modal elicited production study aimed to measure which modals children prefer to use in experimentally controlled modal situations crossing force (possibility, necessity) and flavor (teleological, epistemic). We specifically asked whether children make distinctions by force and flavor in their modal usages by situation, and compared them to adult controls for the same region.

Adult results confirm that our task elicited the kind of modal productions we expected: adults prefer to use possibility modals in possibility conditions, and necessity modals in necessity conditions, for both flavors (noting there is less necessity modal use overall, and two adults used only possibility modals). Adults used different modals for difference forces. This pattern is true in aggregate, and for most individual adult's usage patterns. For flavor, adults tend to use the same modals in possibility conditions (preferring *could* for both teleological and epistemic possibility conditions), but they tend to differentiate by flavor in necessity conditions, preferring *has to* and *should* for teleological necessity, and *must* for epistemic necessity. There is also some flavor-general use in necessity (they use *should* sometimes in epistemic necessity, and *must* sometimes in teleological necessity). Generalizing, adults' preferred modal verbs tend to differentiate the modal meaning space fully by force and partially by flavor. Adult results align best with Map A, supporting Hypothesis A, and partially Hypothesis D (aligning with Map D).

Child results show they prefer to use possibility modals (e.g., *can, could, might…*) over necessity modals (e.g., *must, has to, should…*) in general, regardless of the force condition. Children use possibility modals in necessity contexts (as do adults), but they also use necessity modals in possibility contexts more commonly than adults. Overall, children's modal uses are balanced for each force, but differ by flavor. Only *has to* shows some bias in the adult-like direction for force, as children used it less in teleological possibility than necessity. Children also frequently used future auxiliary verbs: *will* and *would* in all conditions, and *gonna* in teleological contexts;

---

[11] Of the children included in our analysis, 1 produced no modal verbs, yielding a total of 37 child modal verb maps. Of the children who used modal verbs during the experiment, children used an average of 6 modal verbs (standard deviation=4, mode=1, range=1-16).

In contrast, adults only occasionally used *will* (mostly in necessity conditions). Individual child maps show much more variation than adults'; most strikingly, many children used futures. Overall, many children tended to use modal verbs only in one flavor, making it unclear whether they've learned that many English modal verbs can express more than one flavor. Children were also more likely than adults to draw from the wider set of modal language – including belief verbs and adverbs – to express their intended meanings (i.e., they were more likely to diverge somewhat from the grammatical frame we provided, e.g., "Kat <white noise> go down the red road"). However, when children use these other categories, they are in line with children's modal verb usage: adult-like for flavor, but not for force. For example, the epistemic possibility adverb *maybe* was used by some children in epistemic contexts, both in possibility and necessity.

Generalizing, children's preferred modal verbs tend to differentiate the modal meaning space by flavor (the mode modal uses are *can* for teleological flavor and *might* for epistemic flavor), but not by force. Child results align best with Map B, and thus support Hypothesis B, as their modal verb usage shows little to no distinction by force, but clear distinction by flavor. A few modal verbs in child usage (notably *could*, also the future auxiliaries) occurred in all 4 conditions, showing partial support for fully generalist modal use (Map C: force-general and flavor-general). We can conclude that children's modal verb usages show sensitivity to the root vs. epistemic distinction, but not to the possibility vs. necessity distinction. Children in our sample do not show much evidence suggesting they appreciate that hallmark feature of English modal verbs (at least to linguistics (Hacquard, 2011; Kratzer, 1977; i.a.): their flavor-variability.
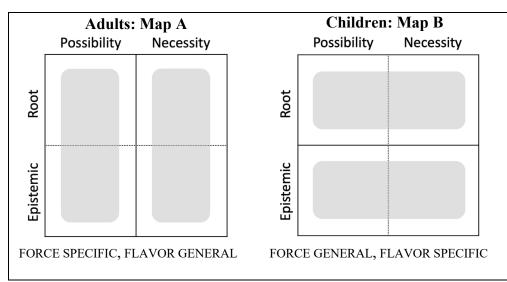

Figure 8. Maps (adapted from Figure 1) that best capture the general aggregate patterns of adult (left) and child (right) modal verb responses, plotted on the 2 x 2 modal meaning space.

Children's responses reflect sensitivity to modal flavor. Whether children possess the conceptual underpinnings necessary to distinguish and track aspects of root and epistemic modalities is under debate. On the one hand, developmental psychology studies argue that infants and toddlers understand various types of root modal reasoning (e.g., desire-intentional, e.g., Tomasello & Carpenter, 2007; obligations, (Cummins, 1996; Dunn, 1988; and goal-motivated behaviors, Woodward, 1998) and aspects of epistemic reasoning (e.g., some awareness of knowledge states of oneself or others, Goupil et al., 2016; Onishi & Baillargeon, 2005; Southgate et al., 2007) in infants and toddlers. However, a body of research also shows that children in our

age group do not distinguish among root modalities that adults do in comprehension tasks (see Shtulman & Phillips, 2018), treating moral (deontic) and physical possibilities as equivalent (i.e., moral or normative rules are treated as inviolable, like physical laws). Our results suggest that children distinguish at least root from epistemic modality. Further research could use our method to probe for further distinctions among root modalities.

Considering force, children are not showing reliable lexical distinctions between possibility and necessity. This could reflect difficultly with reasoning about possibilities at the conceptual level. Much research shows that toddlers and preschoolers behave in non-adult ways on non-linguistic possibility reasoning tasks, notably not preparing for two equally viable outcomes in act-out tasks (Redshaw & Suddendorf, 2016, i.a.) or appearing to randomly omit open possibilities in various others kinds of tasks (Acredolo & Horobin, 1987; Moscati et al., 2017). In an overview paper, Leahy & Carey (2020) suggests that prior to age 4 children are employing a Bayesian guessing strategy to reason about possibility, and are unable to entertain (or actively maintain) open possibilities. Some have argued that this issue with maintaining open possibilities is why children overaccept necessity modals (e.g., *have to, must*) in epistemic necessity scenarios (Ozturk & Papafragou, 2015). If this reasoning issue accounts for modal usage patterns in our study, it would be because children treat open and closed possibilities interchangeably, perhaps mapping all modals they are learning to a general possibility semantics. Alternatively (or in addition), non-adult like force could reflect a word learning challenge that protracts the learning of necessity modals due to the subset problem: necessity entails possibility, so if learners posit modals like *must* or *have to* are possibility modals they will appear right in the vast majority of usage cases and will perhaps require more and better evidence to strengthen their semantics to necessity (see Dieuleveut et al., to appear, for ways in which children could resolve that problem using pragmatic evidence).

Children's responses may tell us something about how they learn from their input. In teleological conditions, children use *can* most often, unlike adults who prefer *could*, *should* and *have to*. In corpus studies of the input to children, we see that *can* is by far the most frequently used modal verb (along with future *will*), with more than 4x as many uses as the next most frequent modal verb (*have to*), and in unnegated form virtually never occurs in epistemic-consistent contexts in corpora (van Dooren et al., to appear). Because of this learning experience, young children are likely confident that *can* is used to express root modalities. Similarly, children use *might* proportionally much more than adults do in our study, and reliably in epistemic conditions, which we suggest may also be because this modal verb is most clearly epistemic in children's learning experience. Corpus studies show that of all the modal verbs, *might* is the most uniformly epistemic in the input to children (van Dooren et al., to appear: 100% or 76% of uses, depending on one's semantic analysis[12]; Cournane, 2021: *might* alone makes up 39% of her whole sample of epistemic uses of all input modal verbs), and is the modal verb American English learning children (like those in our sample), are most likely to use first as an epistemic (Cournane, 2021). The only modal verb that appears to demonstrate any force distinction for children is *has to* – children use this modal verb more often in teleological necessity contexts than possibility ones (albeit at low rates). Notably, adults use this modal verb mostly in teleological necessity too and only negligibly in epistemic necessity. It is the only necessity modal that adults have a clear flavor preference for: *must* is biased towards epistemic but is used several times in teleological as well, and *should* shows the opposite pattern. Perhaps *have to* is easier for children to learn, because it is not as flavor-

---

[12] *Might* is analyzed either as always epistemic (Kratzer, 1991), or as metaphysical when occurring with a future-oriented prejacent (e.g., *That might fall*) (Condoravdi, 2002).

variable in usage (while it is flavor-variable (Hacquard, 2011), it is used very commonly for root necessity in casual speech, see e.g., Tagliamonte & D'Arcy, 2007, and is exceptionally rare for epistemic necessity in child-directed speech: van Dooren et al., to appear, with 6/2401 (0.2%) for UK English, and 4/486 (0.4%) for NA English). Again, this seems to reflect which modal verbs children are hearing in natural situations, which most reliably map to a particular meaning, here a force-flavor combination. Thus, we suggest that our child results may reflect which modal verbs children learn as most uniformly representing a flavor in their input.

Overall, young preschoolers provided far fewer modal verb responses than adults did. This lack of modal use in preschoolers compared to adults could reflect genuine difficulty with modal verbs (and not just with the task). Production studies on young children also show they are producing proportionally fewer utterances with modal verbs than adult speakers in their input; This lower production rate may especially affect epistemic uses of modal verbs (Cournane, 2021; Papafragou, 1998; Shatz et al., 1983; van Dooren et al., 2017; Veselinović & Cournane, 2020), and children gave fewer modal verb responses in our epistemic conditions than our teleological ones. Nearly three times as many children used modal verbs only in teleological conditions than only in epistemic. This is further consistent with the possibility that young children may avoid epistemic modal verb usage, perhaps because they are using alternate strategies (e.g., adverbs like *maybe*, attitude verbs like *think*) or preferring not to mark modality (plain assertions). Of course, with our data we can only see where children can use a modal, not where they can't.

Genuine difficult with modal verbs as an explanation for children's lower modal verb response rates is further suggested by the fact that when children use alternative strategies to express the situation to Mr. Drooly, they are relatively systematic and semantically reasonable. The most common non-modal-verb strategy was simple tensing of the provided bare verb (*go* or *be*), and this was somewhat more common in epistemic condition (e.g., *Nick is in …*). Much literature addresses the differences between plain assertions and epistemic modal assertions, especially with necessity modals (e.g., Lassiter, 2016; Mandelkern, 2019; von Fintel & Gillies, 2010). The felicity of both plain and modalized assertions (e.g., *It's raining*, *It must be raining*) in epistemic necessity contexts is noted in the child literature as perhaps contributing to child difficulties with modals in comprehension tasks (Noveck, 2001; Ozturk & Papafragou, 2015). Basically, *is*-assertions appear to compete with *must*-sentences when the evidence is indirect (as in our epistemic conditions – we cannot see Nick, we have to infer where he could or must be), both sentences are felicitous. Adult participants in Cournane & Pérez-Leroux (2020) did a sentence-preference task for indirect evidence scenarios, and showed a preference for the *must*-marked sentences, but also often selected the plain assertion, while children showed greater preference for the plain assertion in this condition (towards floor for *must*-sentence selection at age 5).

Children who repeated back the sentence frame without making any additions (e.g., *Kat go down…*), effectively responded with bare verbs. The mean age of our participants was only just over 4-years-old, so this may reflect children in our sample being at the tail end of the optional root infinitive stage (Hoekstra & Hyams, 1998; Rizzi, 1993). In English, this stage presents as an overuse of bare verbs (as English does not have a morphological infinitive) (Wexler, 1994). This response was somewhat more common in teleological conditions, which is perhaps not surprising as bare verbs are non-finite and future-oriented (Kat's goals are achieved in the future from the test sentence; while Nick's hiding is ongoing in the present relative to the test sentence). Previous research has suggested that root infinitive uses are modal in child languages (Hoekstra & Hyams, 1998). However, their study on English bare verb uses suggested this pattern didn't hold in English

(on the basis of natural production). Another possibility is that children are using these as imperatives, i.e., telling the participant where to go or where to be. Anecdotally, sometimes children's prosody suggests this (see also Cournane 2014), but without a more careful prosodic analysis (beyond the scope of this paper) this is speculative.

Alternatively, or in addition, children may have had difficulty with our task. This is suggested by our high exclusion rate, which especially affected sampled 3-year-olds. Children this young can be difficult to work with on explicit behavioral tasks. However, it was important to get 3-year-olds' and young 4-year-olds' data, as this is the critical window where corpus studies show English-learning children using the full array of modal verbs, but it is unclear how they contrast their meanings (or not). This window is the bridge between what we've learned so far from natural production (mostly on ages 1 to 3 or 4; e.g., Cournane, 2021; Shatz & Wilcox, 1991; van Dooren et al., 2017; Wells, 1979), and experimental comprehension tasks (mostly on ages 3 or 4 and up; (Cournane & Pérez-Leroux, 2020; Cummins, 1996; Dack & Astington, 2011; Moscati et al., 2017; Noveck, 2001; Noveck et al., 1996; Ozturk & Papafragou, 2015; Shtulman & Phillips, 2018). Working with younger preschoolers with this kind of production task is challenging, but one must keep in mind that the data is "richer" than with typical comprehension measures, as children are not doing a binary or tertiary response task (forced-choice, truth-value judgements, etc.), but selecting language of their own choosing from their own grammar. This work is comparable to natural production in that way, but allows us to get more targeted and contextually supported productions, and it is also comparable to experimental work, as we get multiple responses within and across individuals and groups in controlled situations.

Our aggregate results are comparable to those from Cournane (2014), with 5-year-old Canadian English learning children and adult controls, testing only modal flavor (without controlling for force). As in our study, children were less likely than adults to provide a modal repair that fit the sentence frame provided (84.8 vs. 98.5% of responses): they sometimes inflected the main verb instead (mostly simple present for ability, bare verbs for deontic and teleological, plain assertions with *is* for epistemic). Essentially, the non-modal-verb strategies were very similar to in our study, but rarer. We suggest this is because the children were older ($M = 65$ months compared to our 49 months), and that children likely gradually come to rely more on modal verbs as adults do (in line with corpus study rates of modal usage). Children's modal verb uses in Cournane (2014) were more similar to adults for the root modal conditions she tested (ability, deontic, teleological) than for future and epistemic modality. For epistemic, adults preferred *must* while children preferred *might* (the conditions were more comparable to our necessity condition, so this is in-line with our results). These similarities also suggest that at least some of our patterns generalize to other dialects.

5. Conclusion

In addition to the formidable conceptual challenges modal reasoning involves, modal verbs present learners with complex grammatical form-to-meaning mapping challenges. The English modal system is perhaps particularly difficult: many modal verbs can express the same meaning, and the same modal verb can express different meanings. Moreover, possibility and necessity modals stand in an entailment relation, creating a classic subset problem. For each modal word, children need to figure out both its force and the range of flavors it can express. We've shown that by the preschool years, English-learning children have come to map different modal verbs to different flavors (root vs. epistemic), and most likely clearly distinguish these conceptual differences. However, it is not clear the extent to which they understand that the same modal verbs can express different flavors. Our results further suggest that children struggle with modal force distinctions,

using the same array of modals and future auxiliaries interchangeably for possibility and necessity. This could be due to conceptual challenges with possibility reasoning, or lexical difficulties with attributing the correct semantics to necessity modals due to the subset problem. Finally, we find that children's modal verb preferences differ from those of adults in the same community, for both force and flavor dimensions of modal meaning. These preferences could lead to potential superficial misunderstandings, and should be kept in mind when designing and analyzing modal comprehension studies.

References

Acredolo, C., & Horobin, K. (1987). The child's relational reasoning and tendency toward premature closure. *Developmental Psychology*, *23*(1), 13–21.

Baayen, R. H. 2008. Analyzing linguistic data: A practical introduction to statistics using R. Cambridge, UK: Cambridge University Press.

Berwick, R. C. (1985). The acquisition of syntactic knowledge (Vol. 16). Cambridge, MA: MIT Press.

Bochnak, M. R. (2015). Variable force modality in Washo. *Proceedings of North-East Linguistics Society*, *45*.

Bochnak, M. R., & Matthewson, L. (2015). *Methodologies in semantic fieldwork*. Oxford: Oxford University Press.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9), 341–345.

Brennan, V. (1993). *Root and Epistemic Modal Auxiliary Verbs.* [PhD]. University of Massachusetts.

Byrnes, J. P., & Duff, M. A. (1989). Young children's comprehension of modal expressions. *Cognitive Development*, *4*(4), 369–387.

Cariani Fabrizio and Paolo Santorio. (2018). *Will* done better. *Mind* 127(505)*.*

Condoravdi, Cleo. (2002). Temporal interpretation of modals: Modals for the present and for the past. In D. Beaver, L. Casillas, & B. Clark & S. Kaufmann (eds.), *The Construction of Meaning*, Palo Alto, CA: CSLI Publications

Copley, Bridget. (2009). The Semantics of the Future. Routledge.

Cournane, A. (2014). In Search of L1 Evidence for Diachronic Reanalysis: Mapping Modal Verbs. *Language Acquisition*, *21*(1), 103–117.

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., & Pérez-Leroux, A. T. (2020). Leaving obligations behind: Epistemic incrementation in preschool English. *Language Learning and Development*, *16*(3), 270–291.

Cournane, A., & Tailleur, S. (2020). La production épistémique chez l'enfant francophone: Complexité syntaxique et ordre d'acquisition. *Arborescences : revue d'études françaises*, *10*, 47–72. https://doi.org/10.7202/1081888ar

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Research on the Acquisition of Syntax and Semantics*. Cambridge, MA: 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.

de Villiers, J. G. (2007). The interface of language and Theory of Mind. *Lingua*, *117*(11), 1858–1878.

Deal, A. R. (2011). Modals without scales. *Language*, *87*(3), 559–585.

Dieuleveut, A., van Dooren, A., Cournane, A., & Hacquard, V. (to appear). *Finding the force: How children discern possibility and necessity modals. Natural Language Semantics.*

Dieuleveut, Anouk, van Dooren, Annemarie, Cournane, Ailís, and Valentine Hacquard. (2021). Finding the force: a novel word learning experiment with modals. Proceedings of *Experiments in Linguistic Meaning (ELM) 1.*

Dixon, P. 2008. Models of accuracy in repeated-measures designs. Journal of Memory and Language, 59(4), 447–456.

Dunn, J. (1988). *The beginnings of social understanding*. Harvard University Press.

Fabricius, W. V., Sophian, C., & Wellman, H. M. (1987). Young Children's Sensitivity to Logical Necessity in Their Inferential Search Behavior. *Child Development*, *58*(2), 409–423. https://doi.org/10.2307/1130518

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23–64.

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences*, *113*(13), 3492–3496. https://doi.org/10.1073/pnas.1515129113

Green, M. G. (1979). The developmental relation between cognitive stage and the comprehension of speaker uncertainty. *Child Development*, 666–674.

Gualmini, A., & Schwarz, B. (2009). Solving learnability problems in the acquisition of semantics. *Journal of Semantics*, 26(2), 185-215.

Hacquard, V. (2011). *Modality. Semantics: An international handbook of natural language meaning, ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 1484–1515*. Berlin: De Gruyter Mouton.

Hacquard, V., & Lidz, J. (2016). Children's attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind & Language*.

Hickmann, M., & Bassano, D. (2016). Modality and mood in first language acquisition. In *The Oxford handbook of modality and mood*.

Hirst, W., & Weil, J. (1982). Acquisition of epistemic and deontic meaning of modals. *Journal of Child Language*, *9*(03), 659–666.

Hoekstra, T., & Hyams, N. (1998). Aspects of root infinitives. *Lingua*, *106*(1), 81–112.

Hoffmann, T. (1966). Past tense replacement and the English modal auxiliary system. *Harvard Computation Laboratory Report NSF*, *17*.

Horn, L. R. (1972). *On the Semantic Properties of Logical Operators in English* [PhD]. University of California.

Kaufmann, Stefan 2005. Conditional truth and future reference. *Journal of Semantics* 22(3). Pp. 231–80.

Kissine, Mikhail. 2008. Why 'will' is not a modal. *Natural Language Semantics* 16(2), 129–55.

Klecha, Peter. (2013). Diagnosing modality in predictive expressions. *Journal of Semantics* 31(3), 443–55.

Kratzer, A. (1977). What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, *1*(3), 337–355.

Kratzer, A. (1991). Modality. Semantics: An international handbook of contemporary research, 7, 639-650.

Kratzer, A. (2012). *Modals and Conditionals, New and Revised Perspectives*. Oxford University Press.

Lassiter, D. (2010). Gradable epistemic modals, probability, and scale structure. *Semantics and Linguistic Theory*, *20*, 197–215.

Lassiter, D. (2016). Must, knowledge, and (in) directness. *Natural Language Semantics*, *24*(2), 117–163.

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

Mandelkern, M. (2019). What 'must' adds. *Linguistics and Philosophy*, *42*(3), 225–266. https://doi.org/10.1007/s10988-018-9246-y

Matthewson, L. (2004). On the Methodology of Semantic Fieldwork. *International Journal of American Linguistics*, *70*(4), 369–415.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. 2017. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315.

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

Nauze, F. D. (2008). *Modality in typological perspective*. Institute for Logic, Language and Computation, Universiteit van Amsterdam.

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

Noveck, I. A., Ho, S., & Sera, M. (1996). Children's understanding of epistemic modals. *Journal of Child Language*, *23*(03), 621–643.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258.

Ozturk, O., & Papafragou, A. (2015). The Acquisition of Epistemic Modality: From Semantic Meaning to Pragmatic Interpretation. *Language Learning and Development*, *11*(3), 191–214.

Palmer, F. R. (2001). *Mood and modality*. Cambridge (UK): Cambridge University Press.

Papafragou, A. (1998). The Acquisition of Modality: Implications for Theories of Semantic Representation. *Mind & Language*, *13*(3), 370–399.

Piantadosi, S. T. (2011). Learning and the language of thought, Doctoral dissertation, MIT.

Piéraut-Le Bonniec, G. (1980). *The Development of Modal Reasoning Genesis of Necessity and Possibility Notions*.

Portner, P. (2009). *Modality*. Oxford, UK: Oxford University Press.

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://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, 1-37.

Ramchand, G. (2018). *Situations and syntactic structures*. Cambridge, MA: MIT Press.

Redshaw, J., & Suddendorf, T. (2016). Children's and apes' preparatory responses to two mutually exclusive possibilities. *Current Biology*, *26*(13), 1758–1762.

Rizzi, L. (1993). Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives. *Language Acquisition*, *3*(4), 371–393.

Roberts, I. G. (1985). Agreement parameters and the development of English modal auxiliaries. *Natural Language & Linguistic Theory*, *3*, 21–58.

Ross, J. R. (1967). *Constraints on variables in syntax.* http://eric.ed.gov/?id=ED016965

Rullmann, H., Matthewson, L., & Davis, H. (2008). Modals as distributive indefinites. *Natural Language Semantics*, *16*(4), 317–357.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, *14*(3), 301–321.

Shatz, M., & Wilcox, S. A. (1991). Constraints on the acquisition of English modals. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in Development* (pp. 319–353). Cambridge (UK): Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511983689.010

Shtulman, A., & Phillips, J. (2018). Differentiating "could" from "should": Developmental changes in modal cognition. *Journal of Experimental Child Psychology*, *165*, 161–182. https://doi.org/10.1016/j.jecp.2017.05.012

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–592.

Stephany, U. (1979). Modality. In P. Fletcher & M. Garman (Eds.), *Language acquisition* (2nd 1986, pp. 375–400). Cambridge (UK): Cambridge University Press.

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

Tagliamonte, S. A., & D'Arcy, A. (2007). The modals of obligation/necessity in Canadian perspective. *English World-Wide*, *28*(1), 47–87. https://doi.org/10.1075/eww.28.1.04tag

Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*(1), 121–125.

van Dooren, A., Dieuleveut, A., Cournane, A., & Hacquard, V. (2017). Learning what must and can must and can mean. *Proceedings of the 21st Amsterdam Colloquium*. Amsterdam Colloquium.

van Dooren, A., Tulling, M., Cournane, A., & Hacquard, V. (2019). Lexical Aspect and Modal Flavor in Dutch. *Proceedings from the 42nd Annual Boston University Conference on Language Development (BUCLD 42)*.

Vander Klok, J. (2012). Tense, aspect, and modality in Paciran Javanese. *McGill University: PhD Dissertation*.

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 26*. Michigan Slavic Publications.

Von Fintel, K., & Gillies, A. S. (2010). Must... Stay... Strong! *Natural Language Semantics*, *18*(4), 351–383.

Wells, G. (1979). Learning and using the auxiliary verb in English. *Language Development*, 250–270.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot, D. & Hornstein, N. (Eds*.) Verb movement.* pp.305-350. Cambridge, MA: Cambridge University Press.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. Psychological review, 114(2), 245.