Leaving obligations behind: Epistemic incrementation in preschool English

Ailís Cournane

New York University

Ana Teresa Pérez-Leroux

University of Toronto

Correspondence concerning this paper should be sent to:

Ailís Cournane, Department of Linguistics, New York University, 10 Washington Pl., New York, NY, 10003, USA.

Contact: cournane@nyu.edu

Ana Teresa Pérez-Leroux, Departments of Spanish and Linguistics, University College, Rm F201, University of Toronto, 15 King's College Circle, Toronto ON, M5S 3H7, CANADA.

Contact: at.perez.leroux@utoronto.ca

Abstract

Does language development drive language change? A common account of language change attributes the regularity of certain patterns to children's learning biases. The present study examines these predictions for change-in-progress in the use of *must* in Toronto English. Historically, modal verbs like *must* start with root (deontic) meanings, eventually developing epistemic (probability) meanings in addition. Epistemic uses increase over successive generations, phasing out root uses (INCREMENTATION). The modal becomes unambiguously epistemic and eventually disappears from the language. Such cyclic changes are predictable and common across languages. To explore whether children contribute to incrementation and loss, we tested intuitions about *must* in preschoolers (n=141) and adults (n = 29). In a picture preference task (deontic vs. epistemic) children selected epistemic interpretations of ambiguous sentences (e.g., *Michelle must swim*) at higher rates than adults. Two context-based preference tasks tested children's overall sensitivity to the presence of modals. We found sensitivity in deontic contexts. In epistemic contexts, where *must* is optional and functions like an evidential marker, we found little discrimination, and general avoidance of the modal. These results (epistemic overgeneration, *must*-avoidance) correspond to predictions of the incrementation hypothesis, suggesting children likely play an active role in language change, beyond well-known overregularization processes.

(n= 200)

Keywords: Language Change, Modal Acquisition, Semantic Acquisition, Grammatical Aspect, Sociolinguistic Acquisition, Incrementation, Modality, Change-in-Progress

1. Introduction

Languages change with time. Some changes appear both predictable and directional. What drives those changes? As far back as the 1880s historical linguists speculated that the dynamic nature of child language development underlies regular language change (Paul, 1880, i.a.). More recently, theoreticians have posited that children introduce change by reinterpreting properties of the input (Hale, 1998; Lightfoot, 1979; Roberts & Roussou, 2003; van Gelderen, 2011). This study explores whether children drive change in one well-studied case of regular language change: modal verbs.

Modals allow speakers to express non-actuality. Modal meanings divide into two broad flavors. ROOT modality[1] includes abilities, goals, and obligations, among others. EPISTEMIC modality includes inferences from knowledge or evidence. We focus on *must*, which like many other modals verbs in English and other languages is flavour-variable and can be interpreted as either root (1a) or epistemic (1b).

(1) Penguin must eat a lot of fish …otherwise he will freeze. (Root: Obligation or Goal)
…because he's plump. (Epistemic: Inference)

Interestingly, flavour-variable modals commonly change unidirectionally, from historically older root uses to newer epistemic ones (Bybee, Perkins, & Pagliuca, 1994; Lightfoot, 1979; Traugott, 1989, 2006). The predictability of these diachronic processes begs for inquiry into language transmission and the role of language acquisition. While the standard method in language acquisition is to measure the child against an adult norm (since Brown, 1973), full convergence is not the only possible outcome. Children might change the system as they learn it. Robust, regular diachronic patterns represent an optimal place to investigate the role of children in language change (Cournane, 2014, 2015, 2017).

[1]Root modality (Hoffmann, 1966) includes all non-epistemic meanings that attribute the possibility or necessity of the predicate to the subject of the verbal event. Deontic modality is one subtype of root modality.

Standard methodology implicitly predicates the existence of a stable reference point – a target system of forms and meanings. However, language (and language input) is neither homogenous nor temporally stable (Kroch, 1989; van Gelderen, 2011; Weinreich, Labov, & Herzog, 1968). Recent studies have tackled this issue directly, by examining how variability in the input affects the timetable and outcome of development, including the effect of phonetic variation in the formation of grammatical categories (Miller, 2012; Schmale, Cristià, Seidl, & Johnson, 2010), and to what extent prosodic variability affects children's performance (Sundara, Demuth, & Kuhl, 2011; Ettlinger & Zapf, 2011). Artificial language studies show that child learners extract regular patterns from inconsistent input (Hudson Kam & Newport, 2005). A current line of work investigates how much noise can be tolerated before disrupting rule generalization (Schuler, Yang, & Newport, 2016; Yang, 2016; Hendricks, Miller, & Jackson, 2018). In natural contexts, work on Nicaraguan Sign Language explores how younger learners introduce language systematicity (Cohort 2 and later; e.g., Senghas & Coppola, 2001). A separate literature studies to what extent children are able to mirror sociolinguistic patterns of variability. Some of this work explicitly assumes community-level directional patterns of generational change exist, but that child learners do not generate them (Labov, 2001, 2007; Smith, Durham, & Richards, 2013; i.a.). Finally, usage-based approaches to first language acquisition stress individual differences and non-homogeneity of input, arguing these are an "inconvenient truth" for generative approaches which standardly abstract away from variability (Kidd, Donnelly, & Christiansen, 2018).

Our focus is on the question of what role children might play in generating such variation (as opposed to only experiencing it) and in determining its directionality (in the case of regular cyclic historical changes). A perspective from the study of diachronic syntax holds that children play an active role in language change, as creators and selectors of the target of change (Lightfoot,

1979; Roberts & Roussou, 2003; van Gelderen, 2011). While this is accepted for regularization processes (Kam & Newport, 2009; Fedzechkina et al., 2016), it has not yet been tested for cyclical changes (van Gelderen, 2011), that is, an identified set of recurrent patterns that affect certain words, such as modals (Bybee et al., 1994) or negation (Jespersen, 1917; van Gelderen, 2011). We assume that both children and languages are dynamic systems changing over time, and propose to test whether learner dynamics underlie the dynamics of language change in the case of modal verbs by extracting predictions from what is known from the diachronic modal cycle.

English auxiliary modal verbs (e.g., *can, will, might*) derive from the past tense forms of West Germanic/Old English main verbs. *Must* was first recorded as a content verb with an ability/permission meaning (cf. *be able/allowed to*)[2] (2a). By Middle English it was a flavour-variable auxiliary modal verb (Lightfoot, 1979; Visser, 1963), with primarily root usage (Traugott, 2006) (2b). A gradual shift towards epistemic uses is still in progress, most advanced in some varieties of English such as the one spoken in Toronto (Tagliamonte & D'Arcy, 2007) (2c). As one modal changes, historically new forms take over older uses, like *have to* and *got to* in English, which are currently primarily used with root meanings that were previously associated with *must* (Fischer, 1994; Krug, 2000). The modal cycle is renewed over time with new words. The Toronto variety of English is one where the use of *have to,* over *must,* increases sharply across age spans (preadolescents, adolescents, adults, and seniors); suggesting an advanced stage in the cycle for *must*. Preadolescent girls use *have to* for obligation meanings about 25% more than older women. The usage differences by age group are more extreme for males, namely older men use more of old-fashioned *must* than modern *have to*. The preadolescent and adolescent groups in Tagliamonte

[2] In its history *must* strengthened from a possibility modal to a necessity modal, e.g., permission > obligation, a change not central to our current study (Yanovich, 2013).

and D'Arcy's (2007) study are significantly different from each other, with usage rates varying by age group, a typical indicator of change-in-progress over time in a language[3].

(2) *Must* in diachrony (based on Traugott, 2006; Oxford English Dictionary)

(a) *mote*$_V$, past.*mōste*		(b) *most(e), must*		(c) *must*
<O.E (400-1100CE)	→	M.E. (1100-1500CE)	→	Mod.E. (1500CE-Today)
~be able/permitted to		*obligation, (epistemic)*		*(obligation), epistemic*

Each stage in the cycle represents a distinct process that may reflect child innovation (Figure 1). First, GRAMMATICALIZATION represents how loss of the lexical meaning of a verb with a related sense (e.g., *oblige*, *wish*; 'premodals') yields a functional auxiliary modal (Lightfoot, 1979). This change allows the modal to take scope over the clause, which yields a novel epistemic sense via meaning OVEREXTENSION (possible only once the modal is functional; Hacquard, 2013). The item, now flavour-variable, undergoes INCREMENTATION of epistemic uses, alongside gradual loss of root uses (Tagliamonte & D'Arcy, 2009). Finally, in overlap with the flavor-variable stage, the modal undergoes USAGE REDUCTION, where overall rates of use drop as newer forms join the Modal Cycle (e.g., *have to*) and gain usage. These usage shifts – fewer overall uses with a concomitant increasing proportion of epistemic uses – take place gradually and directionally over many centuries, and are currently active across all dialects of English. During all stages, syntax regulates interpretation (Hacquard, 2013), a point we will return to.

[3] In variationist sociolinguistics, cross-sectional differences between generations are taken as a proxy for ongoing language changes. This is called change in *apparent time*, as contrasted to changes in an individual's usage patterns through their life, which is change in *real time* (Labov, 2001; Tagliamonte & D'Arcy, 2009, i.a.).

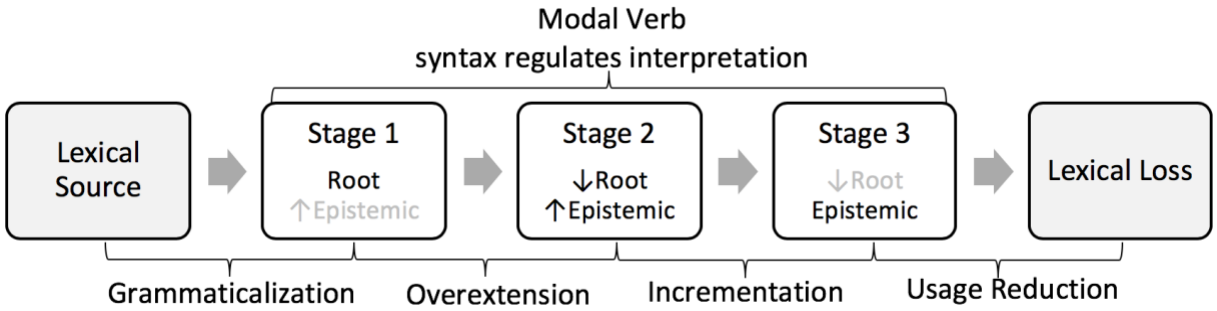


Figure 1. Cyclical stages of modal lexemes. 1. Grammaticalization bleaches the original lexical sense of the source verb and yields a modal verb (Stage 1). 2. An overextension process introduces flavor ambiguity (Stage 2). 3. Epistemic incrementation leads to a reduction in use of root senses (Stage 3). 4. Lexical Loss. At Stage 3, root uses are associated with formal/written registers, and newer modals participating in the cycle crowd out older forms in usage, leading to lexical loss.

Must offers an opportunity to study how competition between word senses (obligation and probability) evolve and to examine the role of children in such changes. To learn the word, children need to navigate two families of meanings (root, epistemic), and to understand the interaction between meaning and sentence context. When accompanied by grammatical aspect marking (*must be sleeping*, *must have slept*), *must* is only epistemic. We consider the case of Toronto English, where *must* is at an advanced stage of epistemic incrementation, to explore the following questions:

i. Do children play a role in furthering epistemic incrementation of modal verbs?

ii. What is the role of syntax in supporting the acquisition of modal verbs?

We first discuss how syntax interacts with modal meanings, and examine the role played by aspect-marking and temporal orientation of the predicate. We then review the previous literature on modal acquisition, and present our three studies testing children's interpretation of *must*.

1.1. How syntax interacts with modal meanings

To understand how children learn to map the meaning of modals, we propose to start by clarifying the relationship between modality and sentence structure, including lexical and grammatical aspect. While the situational context is important for how adults interpret modals (Kratzer, 2012), it may or may not be sufficient to support learning. An alternative is that children may exploit the robust cross-

linguistic correlations between modal syntax and temporal orientation (van Dooren, Dieuleveut, Cournane & Hacquard, 2017; Condoravdi, 2002; Werner, 2006). In general, root modals are only future-oriented, with the event following the time of utterance (i.e. goals and obligations are about the future). Epistemic modals can be past or present-orientated, evaluating the likelihood of events that have occurred or are currently occurring (i.e., uncertainty of the modal is relegated to mental worlds: one can say *x must be happening* when not fully informed about current state of affairs).

Lexical and grammatical aspect impact modal interpretation. In English, perfect aspect marking yields a past temporal orientation where the inference from *must* is about a previous situation (3a). Progressive aspect yields a present temporal orientation where the inference from *must* is about what is happening now (3b). Grammatical aspect markers thus key epistemic interpretations.

(3) a. Doggy must [have eaten his dogfood]$_{\text{perfect}}$ (past-oriented: *root, epistemic)

b. Doggy must [be eating his dogfood]$_{\text{progressive}}$ (present-oriented: *root, epistemic)

Modal meanings are linked to syntactic positions as shown in (4): Root interpretations occur when the modal is above the verb phrase scopes below tense. Epistemic interpretations arise when the modal scopes over both tense and aspect, predicated of the entire proposition ('must be the case that *p*') (Hacquard, 2006; cf. Matthewson, 2012). Bare verb constructions remain ambiguous whereas aspect marking disambiguates the position of the modal.

(4) [Subject [***must***$_{\text{EPISTEMIC}}$ [Tense [**Aspect** [*must*$_{\text{ROOT}}$ [Verb Phrase]]]]]]

Absent grammatical aspect, the lexical aspect of the main verb can also inform temporal orientation and affect the interpretation. Stative verbs (i.e., *be, love*) trigger present-orientation and exclude root meanings (5a). Eventive verbs (i.e., *eat, go*) yield either future-orientation and root interpretations, or present habitual construals with epistemic interpretations (5b). Importantly, perfect and progressive markers also have the effect of changing event predicates into states.

(5) a. She must be in bed. (stative verb: *root, epistemic)

b. She must go to the gym. (eventive verb: root, epistemic)

Temporal orientation, as encoded in lexical and grammatical aspect, could provide regular cues for the learner (van Dooren et al., 2017; van Dooren, Tulling, Cournane, & Hacquard, 2019). Children are sensitive to aspect, both lexical and grammatical from a very early age (Wagner, 2001). Van Dooren et al. (2017) analyzed a sample of 20,528 English child-directed utterances with modals, finding modals occur with grammatical aspect 11% of the time when used epistemically, but less than 1% when used as root. Absent grammatical aspect, modal verbs used with root senses were overwhelmingly paired with eventive verbs (94%), while modals used with epistemic senses appeared with stative verbs (83%). Van Dooren et al. (2019) found even stronger patterns in Dutch. These distributional cues support the viability of a syntactic bootstrapping approach for learning epistemic meanings of modal verbs.

1.2. How children learn modals

Modals are a classic case of hard words whose meanings must be carefully extracted from association with sentence context (Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; van Dooren et al., 2017). They encode the likelihood of non-actual or unconfirmed events or propositions, often making reference to abstract mental state notions such as goals, speaker-knowledge, or socially complex rules and norms. Flavour-variability further complicates the learning scenario (Papafragou, 1998; van Dooren et al., 2017). The standard view of the acquisition of modal verbs attributes a key role to conceptual development (Astington, Harris, & Olson, 1990; Kuczaj & Daly, 1979; Papafragou, 1998; Stephany, 1979, i.a.). These approaches rely on naturalistic data showing that modal verbs emerge at age 2, but only with their root meanings (e.g., ability, intention, permission). Epistemic meanings appear later, around age 3. Sarah (Brown, 1973) shows the typical

pattern of development. Her root modal uses are present from the outset, first with ability uses (6a), then with deontic (6b). First epistemic use is found at age 3 (6c).

(6)	a.	I can ride one. (= a toy horse)	Sarah 2;04,12	Root (Ability)
	b.	I can't do it. (=hurt her mom)	Sarah 2;11,28	Root (Deontic)
	c.	Must be gone. (=missing toy dishes)	Sarah 3;00,27	Epistemic

Some remarks about conceptual approaches are in order. First, the observation of root-before-epistemic holds exclusively for functional modals (Cournane, 2015), which are both polysemous and grammatically complex (Hacquard, 2006, 2013). Adverbs can also convey the same epistemic senses (Kratzer, 2012; Traugott, 2006), and children use adverbs like *maybe* in clear epistemic contexts as early as age 2 (7) (Cournane, 2015; O'Neill & Atance, 2000).

(7) *maybe grandma made this.* (inferring about the origin of a handicraft object)

Violet, 2;03.05 (Providence Corpus, Demuth, Culbertson, & Alter, 2006)

Furthermore, data from implicit methodologies (gaze-fixation patterns) suggest that infants are more conceptually sophisticated than previously considered (e.g., Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007). Recent data from denser child corpora (Manchester Corpus; Theakston, Lieven, Pine, and Rowland, 2001) revises the age of first use of epistemic modal verbs (van Dooren et al., 2017). Epistemic uses are attested at 2, but root uses are much more frequent than the input would predict. For example, in the Manchester Corpus *must* is ~80% epistemic in adult speech but ~80% root in 2-year-old children. This corpus also shows alternating input use of root and epistemic senses at variable rates for other modal lexemes (e.g., *might, may, could*).

Most experimental studies on modal meaning have focused on strength contrasts (e.g., *must* vs. *might*; Hirst & Weil, 1982; Noveck, 2001; Ozturk & Papafragou, 2015), and few studies consider the problem of root vs. epistemic flavor. Unpublished work by Fond (2003) used a picture-choice task to test if 4-year-olds exhibited a deontic bias for *must* and Spanish *deber* (*v.*'must'). Children

listened to bare modal sentences such as "*The dog must eat a lot*", choosing whether to match them to deontic (e.g., a skinny dog who needs to eat) or epistemic scenarios (i.e., a fat dog). English-speaking 4-year-olds were bimodal, between deontic and epistemic bias.

Heizmann's (2006) preliminary study was the first to test the role of grammar in modal interpretation. She tested preschoolers' interpretations of English *must* and German *müssen* (*v.* 'must') in a context-based task. Children were presented with questions like: *Who must be eating a banana*? (progressive morphology in English makes this receive an epistemic reading) or *Who must eat a banana*? (deontic interpretations available). All English preschoolers (aged 3–5) appeared to be able to select epistemic-interpretations with unambiguous sentences. For ambiguous structures, children in both languages exhibited an early preference for deontic readings.

In a sentence preference task, given an epistemic context, 5-year-old children have been shown to often choose what are currently primarily root modals, *have to* and *can* instead of *must* and *might*, respectively (Cournane, 2014). In the same task, adults almost uniformly prefer the established epistemic-biased modals *must* and *might*. This is interpreted as evidence of children showing a unidirectional bias towards extending the root-biased modals to epistemic uses. In root contexts, in contrast, both children and adults equally allow the alternative lexeme (allowing *must* in obligation and *might* for permission roughly 18% of the time). This is evidence that these modals are used ambiguously for both children and adults. In a related sentence repair task that masked the position of the modal (i.e., *Sam <<noise masking>> fall in the water*), children used innovative responses, relying more than adults on quasi-modal alternatives (e.g., *have to* instead of *must, gonna* instead of *will)*. This is predicted under the view that children lead the way with modal language changes (Figure 1).

To summarize, flavour-variable modals occur with root uses as early as age 2, whereas epistemic uses typically emerge after age 3 and are initially low relative to the input (Cournane, 2015; van Dooren et al., 2017). Preschoolers appear able to access both root and epistemic interpretations for modal verbs (Fond, 2003; Heizmann, 2006). These preliminary results suggest children might have some sensitivity to syntactic context (Heizmann, 2006). Finally, there is also preliminary evidence that children's modal selection deviates from adults' in the predicted direction of diachronic change (Cournane, 2014).

1.3. Hypotheses

How do children learn epistemic meanings and what is the role of syntax? The evidence suggests that epistemic reasoning and some forms of epistemic language are present from age 2 but are rare at the outset. Therefore, the conceptual approach must be softened to allow for early capacity for representing epistemic reasoning, along with changes in capacity to access those representations in language use, possibly related to domain general development.

Syntactic bootstrapping approaches hold that children can use distributional information to learn the meanings of difficult words. In the case of modal verbs, the most relevant cues are tense and aspect marking. The literature on modal verb input for English and Dutch shows that the strongest asymmetry in the input is that when used as root, modal verbs usually have eventive verbs in their complement, but when used as epistemic, they have stative complements (either stative verbs, or eventive verbs marked with grammatical aspect). As assumed in the semantic literature, grammatical aspect turns eventive verbs into statives. Thus, it's possible the distribution of aspect might serve as a cue to which senses are possible and preferred for a given modal lexeme.

Additional assumptions about the timing of different developmental biases arise not from the acquisition literature, but from the linguistic study of language variation and change. Contrary

to the common theoretical and normative practice of abstracting away from language variability, work by William Labov and colleagues (Labov, 2001; Weinreich, Labov, & Herzog, 1968; i.a.) has yielded a solid body of evidence on the multi-dimensionality of language variability, which ranges along parameters of geography, gender, social class, and particularly important to our case, age. More often than not, these systematic and measurable differences are imperceptible to speakers themselves (Labov, 2001, 2007), at least at a conscious level. The effect of this is that what is synchronically described "language" is a heterogeneous object containing different stages in language change-in-progress (Kroch, 1989).

It is useful to distinguish between the concepts of representation as opposed to usage. New representations may arise in language, without the loss of older representations (van Gelderen, 2011). These can even coexist in the same morpheme, which then is described as polysemous. Two speakers may have access to the same representations of a form, while allowing it to be used in a different range of constructions, and at different rates (compared to competitors, e.g., *must* vs *have to*). Such contextually narrow distributional differences in usage might impact child learning. It is reasonable to propose that children's experience is likely to contain speakers at different diachronic stages (Figure 1), including individuals who use *must*: (a) variably as root and epistemic (late Stage 2), (b) mostly as epistemic (Stage 3), (c) exclusively as epistemic (late Stage 3), or, (d) rarely, or at all (Lexical Loss Stage, using *have to, got to* instead). Within the ongoing evolution of *must* in Toronto English, which is more advanced than many varieties of English, we predict that input for *must* is reduced to the point that some children will avoid it altogether, even in contexts that call for it in adults. If the existence and directionality of incrementation and lexical loss comes from child learning "errors" and biases, as proposed in the theoretical literature, we expect children to be biased towards the direction predicted in cyclic changes.

On the basis of these assumptions, we evaluate the following hypotheses and conjecture about the course of acquisition of modal meanings. These hypotheses are not contrastive.

H1. EPISTEMIC GROWTH HYPOTHESIS: Child preference for epistemic interpretation will grow with age.

H2. ASPECTUAL BOOTSTRAPPING HYPOTHESIS: If children are using aspect to bootstrap the assignment of epistemic meanings to English modals, children should avoid giving root meaning to aspect-marked modal sentences.

H3. EPISTEMIC INCREMENTATION HYPOTHESIS: We predict a stage where children will overgenerate the epistemic interpretation of *must* + bare verb, when compared to adults.

H4. MODAL DETECTION HYPOTHESIS: Children are aware of the presence of the modal, despite not necessarily matching adult preferences and interpretations.

- USAGE REDUCTION CONJECTURE: Some children will avoid *must* altogether, reflecting the gradual diachronic lexical loss of the verb *must.*

To test these hypotheses we conducted 3 studies. The first study tests the availability of children's interpretation of modal verbs as epistemic along the age range studied (Hypothesis 1 and 3), with and without aspectual marking (Hypothesis 2). The other two studies investigate children's sensitivity to the presence of modal separately for root and epistemic contexts (Hypothesis 4 & usage reduction conjecture).

2. Study 1: Flavor Preference Task

2.1 Methods

2.1.1 Participants

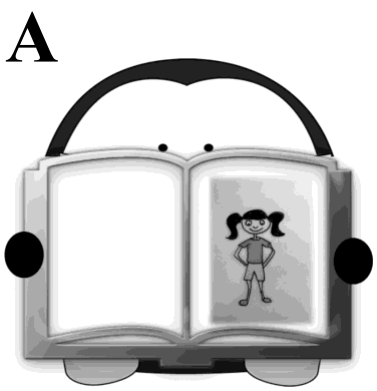
We recruited 54 monolingual English-speaking children, born and raised in Toronto, through personal networks and local daycares. Controls were monolingual adults from Toronto. Children were divided into groups: 3-year-olds (n= 17, range: 36-47 mos, M= 41, SD= 3.5), 4-year-olds (n=18,

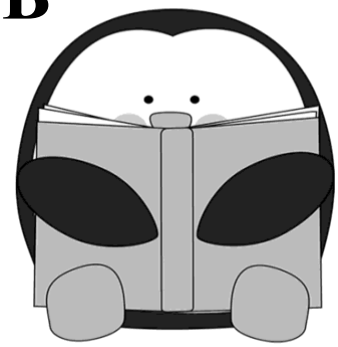
range: 48-59 mos, *M*= 53.3, *SD*= 3.1), and 5–6-year-olds (*n*=19, range: 60-77 mos, *M*= 66.8, *SD*= 4.2). Four other children were excluded, for being bilingual or failing to pass a pretest. Adults (*n*=10, aged 18 to 25) had no exposure to other languages before age 7.

2.1.2. Materials and Procedure

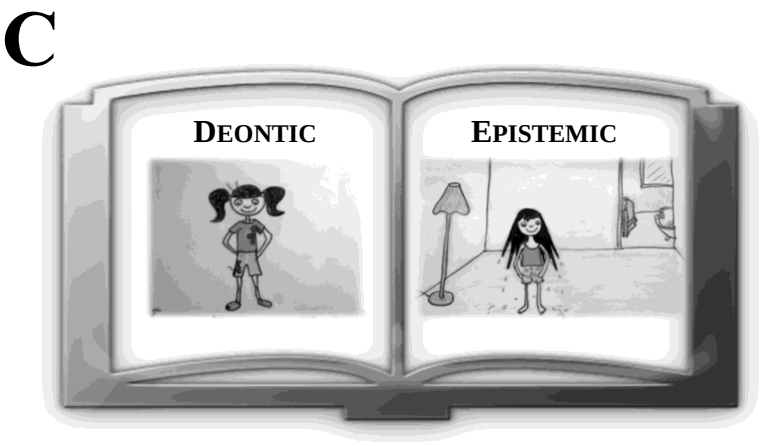
Participants were invited to play a picture-preference game with a character, Penguin, presented on a laptop computer using MATLAB_R2014a with Psychtoolbox (Brainard & Vision, 1997). Responses were recorded using the arrow keys. Penguin's pre-recorded speech has happy affect and natural prosody. Penguin is first shown holding a book (Figure 2-B), with only the cover visible. In the first training item, the book is turned so the picture is presented at the same time as Penguin says, "*Look*! *Bananas*!". Participants are taught to point to the page of the book that "Penguin is talking about". In the next two items, the prompt s heard before the book s turned forward. Data from participants who failed either or both pretest items were excluded. The experimental trials first display an initial picture (Figure 2-A), with Penguin's voice providing a short context (e.g., "*This is Jada. She likes to play in the mud, but she also likes to be clean*"). Penguin turns the book back, flips the page (Figure 2-B), and says the target sentence, and then "*See, look*!" (Figure 2-C) as the book is reoriented towards the participant. The experimenter then prompts, "Penguin said, '*Jada must take a bath*' – which picture was Penguin looking at?".

Set-Up





Test Pictures (DV)



		MODAL ONLY	**MODAL ASPECT**	Total
Test Sentences (IV)	**PERFECT**	4 Sentences: "*Jada **must** take a bath*"	4 Sentences: "*Jada **must** have taken a bath*"	8
	PROGRESSIVE	4 Sentences: "*Scott **must** wear his boots*"	4 Sentences: "*Scott **must** be wearing his boots*"	8
	Total	8	8	16

Figure 2: Flavor Preference Task, Sample Stimuli and Design (2x2). Narration for each image: **(A)** Penguin: *This is Jada. She likes to play in the mud, but she also likes to be clean.* **(B)** [Page turning noise]. Penguin: *Oh! [Test Sentence]*. **(C)** Penguin: *See look!* Experimenter: *Penguin said, [repeat test sentence]. Which picture was penguin looking at?*. Each participant saw 8 stories with MODAL-ONLY sentences and 8 with MODAL-ASPECT sentences, 4 each of which were PROGRESSIVE and PERFECT. The participant was prompted to pick either the DEONTIC or the EPISTEMIC picture.

The contexts were designed to fit either interpretation. Epistemic pictures show indirect evidence for the verbal event, while deontic pictures show scenarios where the verbal event has not yet happened and either an authority figure or salient requirement is present in the context (Cummins,

1996). A sample illustration from a perfect aspect trial is given in Figure 2-C. One side presents Jada covered in mud, needing a bath. This represents a future-oriented event consistent with a deontic interpretation, the preferred response for the bare modal. The other picture shows Jada wearing pajamas, with wet hair and tub visible, illustrating a past-oriented situation, consistent with an epistemic interpretation. Children who seemed distractible were at times asked, "How do you know it was that picture?" to keep them focused on the stories.

2.1.3. Design

The study contained 3 training items before 16 test items and 8 distractors in a randomized order. The first item was always a distractor. The distractors, selected as comparable to experimental items, included 4 doubly modified NPs containing attachment ambiguities (e.g., *The woman with the baby with the hat*), and 4 quantifier items with inverse vs. surface scope (e.g., *Every cat is not singing*; 'not < every' vs. 'every < not'). The test items were balanced for sentence type (modal-only, modal-aspect), and aspect-type (perfect, progressive) was nested within sentence type. All trials contained eventive verbs such as *eat* or *wash*. Progressive stories contrasted present-oriented and future-oriented pictures, while perfect stories contrasted past-oriented and future-oriented pictures. Thus, in modal-only sentences, the target deontic choice was paired to either a present-oriented or past-oriented situation (matching the orientation of the aspect trial it was balanced against).

Every participant saw the same 16 test stories. Which sentence type appeared with which story was randomized and balanced for each participant, as was the position of the pictures (left vs. right). The design is summarized in Figure 2.

2.2. Results

Table 1 reports group means and standard deviation for Epistemic picture choices, by sentence type and aspect condition. As expected, adults overwhelmingly choose epistemic pictures for the modal-

aspect sentences, and prefer root pictures for the modal-only sentences. In contrast, with modal aspect sentences, child age groups show a steady increase in epistemic picture choices, with the biggest increase between 4- and 5-years-old. For modal only sentences, we see an inverted U-shaped pattern, with a comparable jump in epistemic choices in the 5-year-olds. No discrimination between sentence type was observed in any of the child groups.

	MODAL ASPECT			MODAL ONLY		
	PERFECT	PROGRESSIVE	Total	PERFECT	PROGRESSIVE	Total
3-year-olds	2.30(1.05)	1.00(0.87)	3.29(1.50)	1.76(1.35)	1.35(0.49)	3.12(1.36)
4-year-olds	2.00(1.19)	1.89(1.18)	3.89(2.08)	1.89(1.18)	1.61 (1.46)	3.50(2.18)
5-year-olds	3.32(0.58)	2.84(1.21)	6.15(1.50)	3.15(1.01)	2.89(0.99)	6.05(1.78)
Adult	3.80 (0.42)	3.50(0.71)	7.30(0.95)	1.20 (0.49)	1.60 (0.49)	2.80 (2.35)

Table 1. Group means (with standard deviations) of epistemic picture responses by each of the age groups by sentence type and aspect condition. Maximum = 4 (Totals, max = 8).

To test for differences between children and adults in patterns of selection of epistemic pictures (H1: EPISTEMIC GROWTH HYPOTHESIS), we employed the *glmer* function in the statistical package *lme4* in R (R Core Team, 2013). The data was fitted into a generalized linear mixed (logit) model using the maximum likelihood method (Laplace Approximation) (Baayen, 2008; Dixon, 2008; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). Fixed effects included sentence condition (modal-only vs. modal-aspect) and group (children vs. adults), as well as the interaction between group and condition. Participant was entered as a random effect (Epistemic ~ Group * Sentence Condition + (1| Participant)). The model shows highly significant effects of sentence condition (SentenceModalOnly, ß = 2.73, Z = 4.87, $p < .001$), group (GroupChild, ß = -3.47, Z = -6.62, $p < .001$) and highly significant interactions between group and sentence condition (GroupChild:SentenceModalOnly, ß = 3.33, Z = 6.10, $p < 0.001$).

To test whether children use aspect to get the epistemic interpretation (H2: ASPECTUAL

BOOTSTRAPPING HYPOTHESIS), we assess whether they discriminate sentence types for the effect of aspect. We fitted the child data into another similar model (Epistemic ~ Sentence Condition + (1 | Participant)). We see that children as a group do not distinguish sentence types (SentenceModalOnly, ß = 0.14, Z = -0.93, p = 0.35).

To assess whether children select epistemic interpretations for ambiguous sentences beyond adult rates (H3: EPISTEMIC INCREMENTATION HYPOTHESIS), we limit the analysis to modal-only sentences, i.e. *must* with bare verbal complements. This generalized linear mixed (logit) model had selection of epistemic pictures as the dependent variable, and age group as a fixed variable, and participants as random effects. Three-year-olds were set as baseline (Epistemic ~ Age Group + (1 | Participant)). Only 5-year-old children differed from the baseline; showing the predicted overgeneration of epistemic responses (AgeGroupFive, ß = 1.82, Z = 4.56, p < 0.001). The model confirms the observation in Table 2, that 5-year-old children significantly overgenerate the number of epistemic choices for bare modal sentences.

To show developmental shifts in individuals patterns we used a series of scatterplots, with each age group represented in a separate panel, shown in Figure 3. Each participant was classified according to two parameters: the number of trials where they chose the epistemic response for modal-only sentences (y-axis) and the number of trials where they chose the epistemic response for modal-aspect sentences (x-axis). The center axis represents a classification of participants. Participants were classified as 'epistemic avoidant' if they gave less than half of their total responses to both conditions as epistemic (Quadrant c), and 'Epistemic dominant' for more than half epistemic responses overall (Quadrant b). Participants were classified as 'target-oriented' if they chose the epistemic picture more often to the modal-aspect condition, than to the modal-only condition (Quadrant d), and those who had the opposite pattern were labelled 'contrarian' (Quadrant a). Most adults appear in the

target-oriented quadrant. They were close to ceiling for aspect-marked sentences, but ranged in their selections of epistemic pictures for the modal-only sentence (with one adult consistently choosing epistemic pictures). Children rarely discriminated between sentence types. Younger children are primarily 'epistemic avoidant' whereas older children become 'epistemic dominant'.

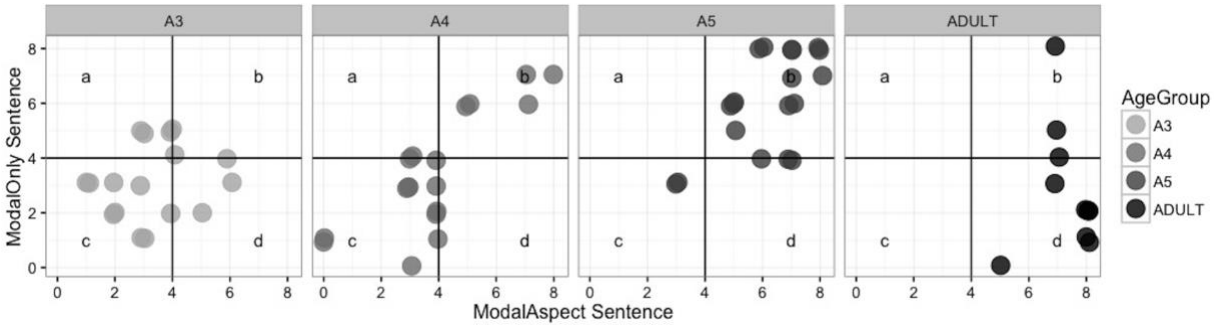


Figure 3. Individual speakers plotted by the number of epistemic picture choices given to the modal-only and modal-aspect sentences. *Quadrant a* is 'contrarian' to target for both sentence conditions. *Quadrant b* is 'epistemic dominant'. *Quadrant c* is 'deontic dominant'. *Quadrant* d is 'target-oriented'.

Children occasionally provided insightful explanations for their responses. A 5-year-old child in (8) illustrates the dominant response pattern by identifying the clues to the inference.

(8) Test Sentence: Modal-only, "Doggy must eat his dogfood" [Child picks Epistemic picture]

a. Experimenter: "How do you know Penguin was looking at that picture?"

b. Child: "His bowl is empty" [points]

2.3. Brief Discussion

Our results show that while adults discriminate between sentence types, children do not, providing no support for the ASPECTUAL BOOTSTRAPPING HYPOTHESIS (H2). Children do not demonstrate understanding that grammatical aspect rules out root interpretations. Instead, they shift from a bias towards deontic interpretations to an overall strong epistemic bias. This is compatible with the EPISTEMIC GROWTH HYPOTHESIS (H1). By age 5 children primarily choose epistemic pictures for modal-aspect sentences at adult-like levels. However, these older children also overgenerate

epistemic interpretations of *must* in bare verb sentences relative to adults, as predicted by the EPISTEMIC INCREMENTATION HYPOTHESIS (H3). A consequence of this increase is that the children fail to demonstrate sensitivity to sentence type.

Two potential issues arise with this task. One, that children may simply undergo age-related shifts in preferences for picture types. The second is the possibility that children who choose the correct picture may do so by relying only on the complement *be eating* vs. *eat* or *have eaten* vs. *eat* without computing the modal interpretation (M. Ippolito, p.c., June, 2014). Aspect markers offer a sufficient cue to temporal interpretation (future-oriented when absent, past or present-oriented with perfect and progressive, respectively). While this may explain children's correct responses, it does little to explain their errors. To test for sensitivity to the modal, we ran two sentence preference tasks. These tasks hold constant the verb phrase to test the effect of the presence or absence of the modal.

3. Study 2: Deontic Detection Task

Study 2 compared simple present sentences (9a) to *must*-sentences (9b). The simple plural has a habitual interpretation with present temporal orientation, while the parallel *must*-sentence has a future temporal orientation, compatible with a deontic interpretation. The temporal situation of the event in (9a) is present-oriented and compatible with a picture where the event is currently taking place (Actual picture). For (9b) the event is future-oriented and compatible with a picture where the event needs to be carried out to satisfy an obligation (Deontic picture).

(9) a. The boys [wash their hands with soap]. No-*must* sentence

b. The boys **must** [wash their hands with soap]. *Must*-sentence

English relies on the present progressive to describe ongoing events, and the simple present has a generic interpretation ('generally does X'). To emphasize the generic construal and make the use of the simple present felicitous, prepositional phrases, like "with soap" were added to the stimuli.

To hold the main verbs constant (i.e., lacking in inflection) in both stimuli types, the subjects of the test sentences were plural and thus did not trigger 3rd singular inflection. This task tests whether children understand that *must* with a bare verb associates with an obligation to carry out in the future.

3.1. Methods

3.1.1. Participants

Participants were 52 monolingual English children, and 10 dialect-matched monolingual English adults (ages 18-22) with similar criteria and recruitment as in Study 1. Children were divided into age groups: 3-year-olds (*n*= 20, range: 38-47 mos, *M*=42.2 mos, *SD*= 3), 4-year-olds (*n*= 15, range: 48-56 mos, *M*= 50.9, *SD*= 2.2), 5–6-year-olds (*n*= 17, range: 60-77mos, *M*= 66.5, *SD*= 4.9). An additional 11 children were excluded for failing the training items (8 of these were 3-year-olds).

3.1.2. Materials and Procedures

Participants were invited to listen to stories told by the experimenter from a storybook. Two puppets, Frog and Shark, controlled by another experimenter, also listened to the stories. Participants were told that Frog and Shark were "silly puppets" who "didn't always pay attention *and* were not always careful with their words". The experimenter invited the puppets to look closely and comment on the picture. Puppets described the scene with sentences differing in the presence of the modal (8). Frog always spoke first to aide memory but sentence type was counterbalanced. To highlight which puppet was talking, that puppet was actively shown moving the mouth, while the other one was lowered to reduce prominence. For each trial, participants were asked to pick the puppet who was paying attention best. In this task, children were systematically prompted on first item and every fourth item, to provide justification for their responses, "How did you know [Shark/Frog] was right?" (Or "Why was [Shark/Frog] right?"). Some children answered consistently, providing spontaneous justification and others gave few or no justifications.

Figure 4 illustrates a typical test trial. The first page introduces two characters (Figure 4-A). The second page (Figure 4-B) shows either an ongoing event (Actual picture) or a cause and a character signaling a command to execute the target action (Deontic picture) (Cummins, 1996). The participant and the puppets had a chance to take in the details of the test picture before prompting the puppets with, "*Look! What do you see, puppets*?". We probed children for justifications of their choices ("How do you know Frog/Shark was right? What did he say?").

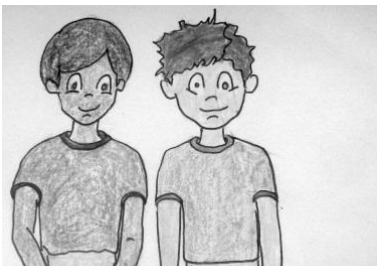
	A 	**B** 	
Set-Up			
Test Pictures (IV)			
	ACTUAL 5 Pictures	**DEONTIC** 5 Pictures	Total 10
Test Sentences (DV)	**NO-MUST** "*The boys wash their hands with soap*"	**MUST** "*The boys **must** wash their hands with soap*"	

Figure 4: Deontic Detection Task, Sample Stimuli and Design (2x1). Narration for each image: **(A)** Experimenter: *These two boys are friends and they went out to play…* **(B)** Experimenter: *Look! What do you see puppets?* Frog: *[Test Sentence 1]*. Shark: *[Test sentence 2]*. Each participant saw 5 stories with ACTUAL pictures and 5 with DEONTIC pictures. The participant was prompted to pick between the puppets, either the one who said the NO-MUST or the MUST sentence.

3.1.3. Design

To verify that participants understood the task a training item showed a picture of a cat and the experimenter says, "*Look*! *What do you see, puppets*?". The puppets looked at the picture, and each

puppet offered a different response, "*A cat*!" vs. "*A dog!"*. The participant was prompted to choose a puppet with "*Which puppet got it right*?". Incorrect or non-responses were followed with direct feedback: the experimenter said "Frog said *"A cat!"* and look [point] it's a cat! So, Frog was right." Two subsequent items were used as pretest. As in the main task a first picture introduced the context "Look! There is a tree" and the target item was presented with a second picture, "A squirrel/bird lives in the tree!". Children who passed both pretests continued to the study. Figure 4 (bottom) summarizes the study design, which consisted of 10 test trials and 5 distractors. Test items were divided per context condition (Actual, Deontic). Trials were presented in one of two random orders.

3.2. Results

Children often answered "both (puppets)", for a total of 65 tokens, evenly distributed across conditions. This is near 13% of all responses. These responses were removed from analysis. Table 2 reports group means. Adult's performance was near categorical, selecting *must*-sentences for Deontic pictures, and avoiding them for Actual pictures. Child categorical answers indicate a slight bias towards the expected distribution, while far from the clear contrast shown by adults.

	ACTUAL		DEONTIC	
	mean	SD	mean	SD
3-year-olds	1.58	(1.57)	2.15	(1.35)
4-year-olds	1.38	(0.96)	1.92	(1.32)
5-year-olds	2.35	(1.58)	3.18	(1.68)
Adult	0.10	(0.32)	4.80	(0.42)

Table 2. Mean (with standard deviations in parenthesis) of 'must' responses given by age groups to picture scenarios (Actual vs. Deontic) and sentence types. (Maximum per cell=5)

To test whether children reliably discriminate for the presence of *must* given the temporal orientation of the scenario (Actual vs. Deontic) (H4: MODAL DETECTION HYPOTHESIS), their data were fitted to a generalized linear mixed (logit) model, similar to those in Study 1, with selection of *must*-

sentence as dependent variable, picture condition as a fixed effect and participant as a random effect (Must ~ Picture + (1| Participant)). We find a significant effect of condition (PictureActual, ß = -0.62, Z value = -3.01, p = 0.003), with children choosing more *must*-sentences for deontic pictures.

Figure 5 plots individual performance. Each participant is represented according to the number of trials where the participant chose the *must*-sentence for Actual pictures (y-axis), and for Deontic pictures (x-axis), following a similar strategy to the one employed in Study 1. Participants were classified thus as 'target-oriented' (quadrant d), '*must*-avoidant' (quadrant c), '*must*-dominant' (quadrant b) and 'contrarian' (quadrant a). Adults are tightly clustered at target, 4-year-olds avoid *must* for Actual pictures, and 5-year-olds are evenly divided between avoiding *must*, target discrimination, and preferring *must* no matter what.

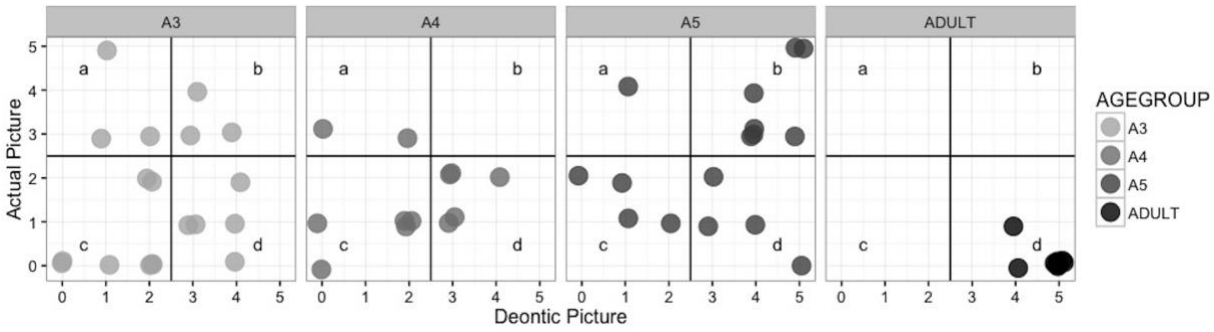


Figure 5. Individual speakers plotted by the number of *must* responses given to the deontic and actual scenarios. *Quadrant a* is 'contrarian' to target for both sentence conditions. *Quadrant b* is '*must*-dominant'. *Quadrant c* is '*must*-avoidant'. *Quadrant* d is 'target-oriented'. N.B. Children who chose "both" on most or all items are not plotted here (n =6).

In this task, children frequently offered relevant explanations in response to our prompt to justify the choice of puppets. These were particularly frequent for the Deontic picture conditions (at the ratio 3:2). These justifications provide qualitative evidence to support that children understand deontic *must.* Responses included some kind of obligation encoding (e.g., repeating *must,* or recasting with other replacing modals (10a,b), negation (10c), imperative structure and prosody (10d), or generic *you* as the subject of the proscribed behaviour (10e), given more for Deontic

pictures (at the ratio 3:1). Some obligation-encoding responses were also given to the Actual pictures (10c), and non-obligation responses, more common for Actual pictures (at the ratio 2:1), included present-oriented descriptions like in (11). Some responses explicitly referenced future-oriented meaning (12a) or translated (12b,c).

(10) a. They must draw on paper but they're drawing on the wall. (Age 5, Deontic, *must*)

b. They need to wear dancing shoes. (Age 5, Actual, *must*)

c. They should not draw on concrete. (Age 5, Deontic, no-*must*)

d. Color on the paper! (Age 3, Deontic, *must*)

e. You wear slippers in ballet class. (Age 3, Deontic, *must*)

(11) a. They're wearing their boots. (Age 5, Actual, *must*)

b. The babies are drinking. (Age 5, Actual, no-*must*)

(12) a. Not different because *must* means maybe they will. (Age 5, Actual, both)

b. *Must* means *have to* (Age 4, Deontic, *must*)

c. They must take a bath, [rephrases]they have to take a bath (Age 3, Deontic, *must*)

3.3. Brief Discussion

Children's choice of *must* for the deontic pictures is reliable but not strong. Non-target responses were given in both directions, and many children also chose both puppets. Qualitative evidence is compatible with the inference that despite the highly variable performance, children appear to understand that *must* associates with obligation, and children show MODAL DETECTION (H4).

We can speculate as to some linguistic factors which may obscure response accuracy. The simple present in English is generic. Children may understand that if one *must* do something one will likely do it. This is commonly known as an actuality inference – if you are obliged to then you do (Traugott & Dasher, 2002). Children who incorrectly chose *must* sentences in the Actual picture

condition, may be choosing not a situation description but a characterization of known norms (Pérez-Leroux, 2016). If they understand its generic/habitual meaning (Pérez-Leroux, Munn & Schmitt, & DeIrish, 2004), they will know the present can be used to express normative behaviours, which are related to obligations ("We always wash our hands before supper")[4].

The sentence stimuli were selected to fix the choice between what must be done, and what is generally done. If these children understood the task as focusing on normative behaviour, rather than as a goodness of fit to the picture, then both sentence types are appropriate This may underlie the prevalence of "both" responses. "Both"-choosers (10/52 children) came from all three child groups. Despite such challenges, children statistically discriminate for sentence types.

4. Study 3: Epistemic Detection Task

Study 3 compared aspect-marked sentences (13a), either with perfect or progressive marking, to counterparts with *must* (13b). The temporal situation of the event is held constant and the modal introduces a contrast in certainty. The *must*-sentences are licensed only when inference is involved rather than direct evidence (Lassiter, 2016; cf. von Fintel & Gillies, 2010).

(13) a. Doggy [has eaten his dogfood]PERFECT — No-*must* sentence

b. Doggy **must** [have eaten his dogfood]PERFECT — *Must* sentence

This task tests understanding that *must* with aspect-marking associates with inferences from evidence about a past or present event. It also serves to test the last stage of the Modal Cycle (Figure 1), where even the more historically innovative epistemic uses of flavour-variable modals start to be lost. In some languages, indirect evidence is optionally marked by a modal (English), while in other

[4] Closer examination of *must*-dominant 5-year-olds suggest the task primed these children for *must*-choices: in their qualitative responses they regularly used *must* deontically (and sometimes *need to*, *have to*), mostly with Deontic pictures. For example, to a picture of girls drawing on the wall, "They must draw on paper" and "They must because they weren't drawing on paper". For Actual pictures, they referenced the obligation associated with the norm (for girls at ballet class: "They need to wear ballet shoes"). Future studies could be between-subjects to alleviate this possibility.

languages it is obligatorily marked by an evidential morpheme (Turkish; Ünal & Papafragou, 2016). To an inferential scenario we can say "Doggy must have eaten his dogfood" or "Doggy has eaten his dogfood". This overlap in felicity, and the optionality of indirect evidence marking, may delay epistemic modal comprehension, as noted in studies on modal strength (Noveck, 2001; Ozturk & Papafragou, 2015).

4.1. Methods

4.1.1. Participants

Participants were 35 monolingual English children, and 9 dialect-matched adults (aged 19 – 30), with similar criteria and recruitment as in Study 1. Children were categorized in age groups: 3-year-olds (n= 9, range: 37-44 mos, M= 40.4, SD= 1.9), 4-year-olds (n= 11, range: 48-59 mos, M= 53, SD= 4.4), 5-year-olds (n= 15, range: 60-71mos, M= 65.2, SD= 3.8). Two children were excluded for being bilingual, 3 for failing either or both of the pretest items and 4 for only choosing the shark puppet.

4.1.2. Materials, Procedures, and Design

As in Study 2, we use a sentence-preference task. Figure 6 illustrates the method. The first picture sets up the context (Figure 6-A). The second picture determines the evidential status as direct (Actual) or indirect (Epistemic) (Figure 6-B). Actual and Epistemic pictures from the same story are provided in Figure 6. The puppets (Frog and Shark), each say a test sentence (with and without *must*), and the participant is prompted to pick between the puppets. Like in Study 2, children were systematically prompted on first item and every fourth item, to provide justification for their responses. Fewer children answered consistently and provided spontaneous justification in this task than in Study 2.

Items included 1 training and 2 pretest items, the same as in Study 2, followed by 16 test items and 8 distractors in one random order. The sixteen test trials were equally divided by context type (Actual, Epistemic) and aspect (Perfect, Progressive). The response variable was presence of

modal (*must*, no-*must*). Within each set, half were assigned to each context condition. Distractors (n= 8) were divided between tests of temporal adverbs or multiple nominal embedding and were similar in length and complexity to the test items. Responses were manually registered on a score sheet. Design is summarized in Figure 6 (bottom).

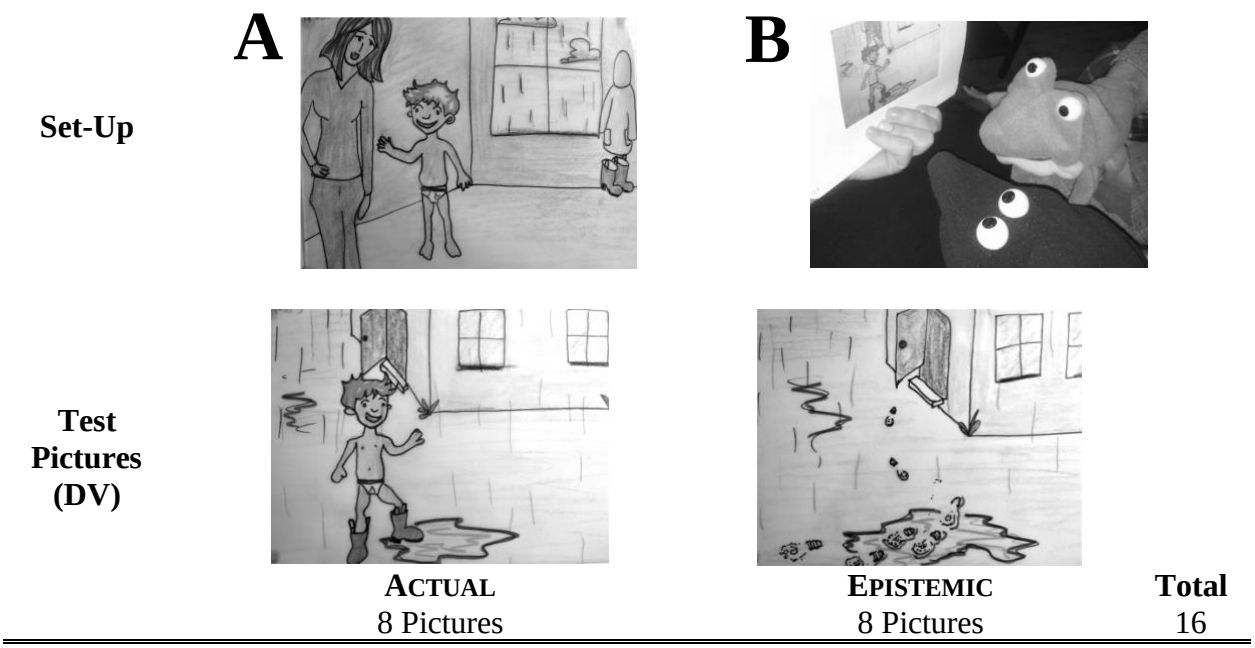


Figure 6: Epistemic Detection Task, Sample Stimuli and Design (2x2). Narration for each image: **(A)** Experimenter: *This is Scott. He wants to play in the rain...but he doesn't wanna wear clothes!* **(B)** Experimenter: *Look! What do you see puppets?* Frog: *[Test Sentence 1]*. Shark: *[Test sentence 2]*. Each participant saw 8 stories with ACTUAL pictures and 8 with EPISTEMIC pictures, 4 each of which were paired with PROGRESSIVE and 4 with PERFECT test sentences.. The participant was prompted to pick between the puppets, either one who said the NO-MUST or the MUST sentence.

4.2. Results

Table 3 reports means and standard deviation for *must*-sentence responses by picture condition by Aspect. We excluded from analyses "both" responses given by one child (age 3) to 5 trials, and one no-answer from a 5-year-old. There is no evidence that aspect condition affects the responses for

either group. Adults responded as expected, with near-categorical choice of no-*must* sentences with Actual contexts (direct evidence), and choice of *must*-sentences more often than no-*must* sentences with Epistemic contexts (indirect evidence), preferring to encode uncertainty. Three-year-olds' *must*-sentence choices are close to chance for both picture conditions. The older children, while continuing to display no sensitivity to picture condition, avoid sentences with *must*.

	ACTUAL			EPISTEMIC		
	PERFECT	PROGRESSIVE	Total	PERFECT	PROGRESSIVE	Total
3-year-olds	1.78(1.09)	2.06(0.81)	3.78 (1.48)	1.89(1.17)	1.78(0.67)	3.67 (1.50)
4-year-olds	0.75(0.87)	0.92(1.16)	1.55 (1.81)	1.25(0.97)	1.17(1.19)	2.27 (2.05)
5-year-olds	0.79(0.80)	0.86(1.17)	1.73 (1.71)	0.86(0.95)	0.93(0.92)	1.93 (1.83)
Adult	0.22(0.44)	0.00(0.00)	0.22 (0.44)	2.44(1.42)	2.56(1.13)	5.00 (2.00)

Table 3. Group means (with standard deviations in parentheses) of 'must' responses given by age group to picture scenario (Actual vs. Epistemic) by aspect condition (Perfect vs. Progressive). Maximum per cell = 4 (for totals, maximum = 8).

To test our MODAL DETECTION HYPOTHESIS (H4) and for differences between adult and child groups in response patterns the data was fitted to a generalized mixed effect model (Must ~ Picture * Group + (1 | Participant)), comparable to previous models. The dependent variable was selection of *must*-sentence, with picture as a fixed effect and participant as a random effect. The model shows highly significant effects of picture condition (PictureEpistemic, ß = 4.68, Z = 5.61, $p < 0.001$), group (GroupChild, ß = 2.77, Z = 3.09, p = 0.002) and highly significant interactions between group and picture condition (PictureEpistemic:GroupChild, ß = -0.62, Z = -3.01, $p < 0.001$). A second model was fitted to the children's data alone (Must ~ Picture + (1 | Participant)). The results confirm the absence of a picture condition effect in children's choice of sentences (PictureEpistemic, ß = 0.24, Z = 1.18, p = 0.24).

As done previously, we plotted the number of trials where the participant chose the *must*-sentence for Actual pictures (y-axis), and Epistemic pictures (x-axis) to classify individual

participants as 'target-oriented' (Quadrant d), '*must*-avoidant' (Quadrant c), '*must*-dominant' (Quadrant b), and 'contrarian' (Quadrant a) (Figure 7). Adults were mostly target, showing variability in the epistemic condition, but consistently avoided modal sentences in the actual condition. Children start by behaving randomly, and become primarily *must*-avoidant by age 5.

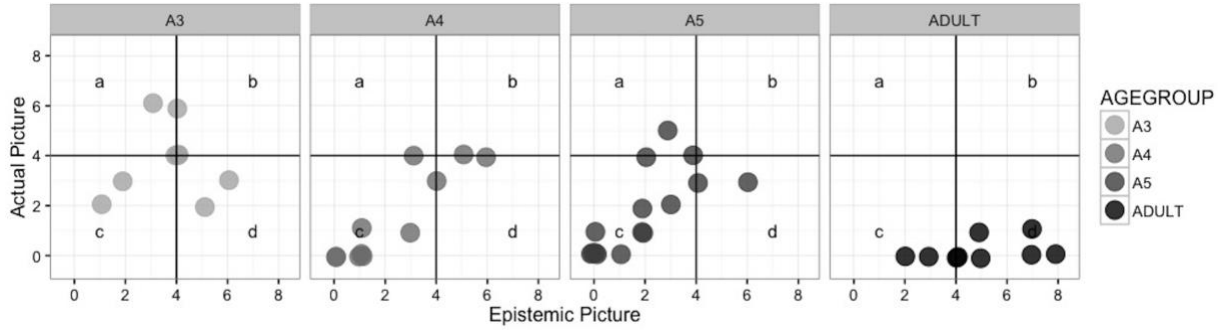


Figure 7. Individual speakers plotted by the number of *must* responses given to the epistemic and actual scenarios. *Quadrant a* is 'contrarian' to expected for both sentence conditions. *Quadrant b* is '*must*-dominant'. *Quadrant c* is '*must*-avoidant'. *Quadrant* d is 'target-oriented'.

This task elicited few informative justifications from children. One 5-year-old who chose only no-*must* sentences was clearly able to justify her choices. She explained to the experimenters that the puppet who used the *must*-sentence was "wrong" because "he said *might*" and sometimes would then point to the picture saying, "but it's not *might* because, look, there's his boots" (Actual picture) or "but the rabbit *did* jump over the fence. The rabbit is gone" (epistemic picture).

4.3. Brief Discussion

Adults behaved as expected, choosing *must*-sentences with epistemic pictures but not with actual pictures. The youngest children are at chance, while older children either remain at chance or avoid *must*. Children show no discrimination by picture, unlike adults, but differ from adults in choosing significantly fewer *must*-sentences in the Epistemic condition, showing MODAL DETECTION (H4) despite non-adult patterns. Many older children avoid *must*-sentences completely or near completely, a pattern consistent with our USAGE REDUCTION CONJECTURE.

Given children's established difficulty with indirect evidentials (Ozturk & Papafragou, 2016), it is not surprising that performance was low. It is possible that children detect the modal, but they cannot evaluate indirect evidence relative to the puppet's perspective (Ünal & Papafragou, 2016). Children may be unable to tease apart epistemic modality (likelihood of *p*) from evidentiality (speaker evidence for *p*).

As most of the older children actively avoid *must* sentences, rather than randomly choosing puppets, it's unlikely that these children are simply 'ignoring' the modal. If these children understand *must* as a possibility operator (i.e., *It must be raining* as a weaker statement than *It is raining*), and cannot yet evaluate indirect evidence relative to another speaker, they may perceive the puppet who says the *must*-sentence as saying something weaker (Lassiter, 2016; cf. von Fintel & Gillies, 2010), which fits neither scenario. Further study is needed to sort out these important questions.

5. General Discussion

The three studies tested children's comprehension of the English flavour-variable modal verb *must*. Study 1 confirmed that child access to epistemic meanings increases with age (EPISTEMIC GROWTH HYPOTHESIS). Age group was significant in both statistical analyses (all responses as well as the subset of modal-only sentences), with older groups choosing epistemic responses more frequently. The full model shows an interaction between sentence type and age group, meaning that adults but not children attend to aspectual cues in choosing their modal interpretation (ASPECTUAL BOOTSTRAPPING HYPOTHESIS). Finally, our results confirm the existence of a stage where children overgenerate epistemic interpretations for sentence types for which adults demonstrate a root bias (EPISTEMIC INCREMENTATION HYPOTHESIS).

Two detection tasks (Studies 2 & 3) probed for sensitivity to the contrast between sentences with and without *must*, and both demonstrated that children treat modal sentences differently from

unmodalized counterparts (MODAL DETECTION HYPOTHESIS). In Study 2, children as a group showed weak but reliable preference for *must*-sentences with Deontic pictures, in the adult-like direction. This observation is fundamental to interpret the epistemic bias found in Study 1, as it shows that older children do not treat *must* as uniformly epistemic but understand its meaning is mediated by syntax. In Study 3 children did not differentiate sentence selection by picture type, but showed *must*-avoidance in older groups (USAGE REDUCTION).

Study 1 results were consistent with 3-year-olds having a weak root bias, even for epistemic aspect-marked sentences. Increased rate of epistemic modal verbs is consistent with the corpus literature (van Dooren et al, 2017; Papafragou, 1998). By age 5, we identify strong epistemic bias, even for modal-only sentences. These results support the incrementation hypothesis: children differ from adults in the direction predicted by historical linguistic theory and by current trends in variation across various age cohorts (Tagliamonte & D'Arcy, 2007). This is the first piece of semantic evidence that children are ahead of adults in the direction predicted by change theories. The patterns of individual variability we observed fit with data on ongoing language change dynamics.

From Study 2 we conclude that children are overall aware of the deontic use of *must* with bare verb complements, showing awareness that future-oriented bare verbs key modal verbs to their root meanings. Like Fond (2003), our evidence supports early polysemy (also Heizmann 2006). In Study 3, the younger children were at chance but the majority of the older children avoided choosing the *must*-sentences, sometimes categorically. This can be interpreted as evidence of usage reduction in the modal, and as consistent with eventual historical lexical loss, the ultimate fate of words participating in cyclic change which get replaced by renewing words (Figure 1).

Why did we observe overall sensitivity to the deontic contribution of *must* in Study 2, but primarily *must*-avoidance in Study 3? Two factors are likely involved. Epistemic meaning are related

to evidence sources, and children are known to have difficulties with grammatical markers of indirect evidential meanings (Ünal & Papafragou, 201). In the task used in Study 3, evidence for the epistemic modal inference is the only difference between picture conditions. The comparison between epistemic *must* and no-modal made the task about the speaker choice to mark indirect evidence or not, because epistemic marking is not obligatory (von Fintel & Gillies, 2010). In languages that obligatorily encode indirect evidence in the grammar (e.g., Turkish) children show adultlike production before adultlike comprehension (Ünal & Papafragou, 2016). In comprehension, those same children overselect direct evidential markers in environments adults select indirect evidentials. With optional English epistemic *must*, perhaps similar difficulties manifest as avoidance. Natural loss of modals in language evolution (at least for languages that do not have grammaticalized evidential markers) may be linked to the increase in epistemic uses: the more epistemic the grammatical marker, the more optional it becomes for speakers.

How can we reconcile results from Study 1, where 5-year-olds strongly prefer epistemic *must*, with those from Study 3, where the same age group in the same speech community avoided epistemic *must*? These two studies tap into different dimensions of the problem: Study 1 addressed what readings are available and preferred for *must* sentences, whereas Study 3 determines, given a particular context, what sentence type was preferred. In short, when the alternatives are root vs. epistemic readings of the same sentence, children prefer epistemic, but when the alternatives are marking epistemicity vs. not marking it, children tend to avoid marking it. This pattern is fully consistent with predictions from the Modal Cycle (Figure 1). Directional change at the latter stages of the Modal Cycle has two phases, first proportional shift towards more epistemic over root uses (Study 1), and second, raw drop in usage rates consistent with gradual loss of the modal (Study 3). These are both happening to individual modals like *must* in ongoing modal change, and individuals

in the same speech community can exhibit different preferences. As one undergraduate at New York University noted about epistemic *must*, e.g., "I know what it means and hear it, but I don't think I ever say that." (D. Halpern, p.c., September, 2017).

Showing that 5-year-olds point towards the direction of change is not an explanation for why they do it. Comparative studies of modal verbs show that modal interpretations are linked to syntactic positions in the following way: when root these modals are represented below tense and aspect, when epistemic they are represented above tense and aspect. In English, when grammatical aspect is present the modal visibly precedes the aspectual marker, offering a potential cue to the epistemic interpretation, which takes scope over aspect. Bare verb constructions, in contrast, are ambiguous as to the scope of *must.* As the epistemic meanings of flavour-variable modals are learned later than root meanings, the epistemic bias may be triggered by the discovery that *must* may occupy a structurally high modal position (Cournane, 2015). Since children do not demonstrate sensitivity to Aspect, this proposal remains speculative.

What other cues do children have for learning whether a modal is epistemic or root? Lexical class of verbs (statives vs. eventives) might prove a strong possibility. Lexical aspect strongly correlates to early tense use, and probably provides a cue for children as to the meaning of tensed forms (Wagner, 2001). Corpus data from English and Dutch (van Dooren et al., 2017, 2019) supports that possibility.

Our effects are small, but that is no surprise. Child divergence is expected to be small, as language change does not take place abruptly but incrementally, attested as small changes in usage over time. This brings up another question: Since children participate in incrementation in a way that is empirically detectable, why does language change not progress more rapidly? One possibility is that once in the school years children will have access to more conservative input with deontic

must, as associated with formal registers. This is part of a well-understood dynamic of retention of historically conservative grammatical features, which possibly plays a role in maintaining the more archaic root meanings of *must*, which are heavily under competition pressure from innovative *have (got) to* in informal language (Tagliamonte & D'Arcy, 2007).

What about overall input biases towards epistemic use? Spoken language in Toronto (Tagliamonte & D'Arcy, 2007) and Manchester, UK (Child Directed Speech; van Dooren et al., 2017) shows a preference for epistemic uses of *must* (with statives, including grammatical aspect). It is possible that as children get older they grow more sensitive to this and input bias (unmediated by syntactic factors) leads to overgeneration of epistemic choices in Study 1. Be that as it may, this leaves the question open as to why speakers become more epistemic-biased for *must* in the first place? Why does change occur unidirectionally in the Modal Cycle and not in the other direction?

In this paper we have provided two critical points of evidence for the child incrementor hypothesis: showing child biases are in the *same direction* as known cyclic changes, and showing children *overshoot* their input. Without the latter, child development could be said to parallel or recapitulate change patterns, but wouldn't contribute to why changes are directional in the first place.

Appendices: Materials for all tasks available at: https://wp.nyu.edu/cournane/papers/materials/

References

Astington, J. W., Harris, P. L., & Olson, D. R. (1990). Developing theories of mind. Cambridge, UK: Cambridge University Press.

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

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. Spatial vision, 10, 433-436.

Brown, R. W. (1973). A first language: The early stages. Cambridge, MA.: Harvard University Press.

Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). The evolution of grammar: Tense, aspect, and modality in the languages of the world. Chicago, IL: University of Chicago Press.

Condoravdi, C. (2002). Temporal Interpretations of modals. In D. Beaver, S. Kaufman, B. Clark (eds.), Stanford Papers in Semantics. Palo Alto, CA: CSLI Publications.

Cournane, A. (2014). In search of L1 evidence for diachronic reanalysis: mapping modal verbs. Language Acquisition, 21(1), 103-117.

Cournane, A. (2015). Modal development: input-divergent L1 acquisition in the direction of diachronic reanalysis (Doctoral Thesis), University of Toronto, Toronto, Canada.

Cournane, A. (2017). In defense of the child innovator. In E. Mathieu & R. Truswell (Eds.), Micro-change and Macro-change in Diachronic Syntax (pp. 10–23). Oxford, UK: Oxford University Press.

Cummins, D. D. (1996). Evidence of deontic reasoning in 3-and 4-year-old children. Memory & Cognition, 24(6), 823-829.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of English. Language and Speech, 49(2), 137–173.

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

Ettlinger, M., & Zapf, J. (2011). The Role of Phonology in Children's Acquisition of the Plural. Language Acquisition, 18(4), 294–313.

Fischer, O. (1994). The development of quasi-auxiliaries in English and changes in word order. Neophilologus, 78(1), 137–164.

Fond, M. J. (2003). Deontic and epistemic modal expression: Theory and acquisition in English and Spanish. (ms.) Smith College, Northampton, MA, USA.

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

Hacquard, V. (2006). Aspects of modality (Doctoral Thesis). Massachusetts Institute of Technology, Cambridge, MA, USA.

Hacquard, V. (2013). The grammatical category of modality. In Proceedings of the 19th Amsterdam Colloquium. Amsterdam, The Netherlands, December 18-20, 2013.

Hale, M. (1998). Diachronic syntax. Syntax, 1(1), 1–18.

Heizmann, T. (2006). Acquisition of Deontic and Epistemic Readings of must and müssen. In Tanja Heizmann (ed.), University of Massachusetts Occasional Papers in Linguistics 34. Amherst, MA: GLSA, University of Massachusetts.

Hendricks, A. E., Miller, K., & Jackson, C. N. (2018). Regularizing Unpredictable Variation: Evidence from a Natural Language Setting. Language Learning and Development, 14(1), 42–60.

Hirst, W., & Weil, J. (1982). Acquisition of epistemic and deontic meaning of modals. Journal of Child Language, 9(3), 659-666.

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

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. Language Learning and Development, 1(2), 151–195.

Jespersen, O. (1917). Negation in English and other languages, volume I, 5 of Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser. Copenhagen: Høst & Søn.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. Trends in Cognitive Sciences, 22(2), 154–169.

Kratzer, A. (2012). Modals and conditionals: New and revised perspectives. Oxford, UK, Oxford University Press.

Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. Language Variation and Change, 1(3), 199–244.

Krug, M. G. (2000). Emerging English modals: A corpus-based study of grammaticalization. Berlin & New York, NY: Walter de Gruyter.

Kuczaj, S. A., & Daly, M. J. (1979). The development of hypothetical reference in the speech of young children. Journal of Child Language, 6(03), 563–579.

Labov, W. (2001). Principles of linguistic change. Volume 2: Social factors. Oxford, UK: Blackwell Publishing.

Labov, W. (2007). Transmission and diffusion. Language, 83(2), 344-387.

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

Lightfoot, D. W. (1979). Principles of diachronic syntax. Cambridge, UK: Cambridge University Press.

Matthewson, L. (2012). On the (non-) future orientation of modals. Proceedings of Sinn und Bedeutung, 16, 431–446.

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.

Miller, K. (2012). Not all children agree: Acquisition of agreement when the input is variable. Language Learning and Development, 8(3), 255–277.

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

O'Neill, D. K., & Atance, C. M. (2000). "Maybe my daddy give me a big piano": The development of children's use of modals to express uncertainty. First Language, 20(58), 029–52.

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.

Ozturk, O., & Papafragou, A. (2016). The acquisition of evidentiality and source monitoring. Language Learning and Development, 12(2), 199–230.

Papafragou, A. (1998). The acquisition of modality: Implications for theories of semantic representation. Mind & language, 13(3), 370-399.

Paul, H. (1880). Prinzipien der Sprachgeschichte [Principles of the history of language]. Tübingen, Germany: Max Niemeyer.

Pérez-Leroux, A. T., Munn, A., Schmitt, C., & DeIrish, M. (2004). Learning definite determiners: Genericity and definiteness in English and Spanish. Supplement to the Proceedings from the 28th Annual Boston University Conference on Language Development (BUCLD 28).

Pérez-Leroux, A. T. (2016). The expression of genericity in child language. In Lidz, J., Snyder, W., & Pater, J (eds.). The Oxford Handbook of Developmental Linguistics (pp.565-586). Oxford, UK: Oxford University Press.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Roberts, I. G., & Roussou, A. (2003). Syntactic change: A minimalist approach to grammaticalization. Cambridge, UK: Cambridge University Press.

Schmale, R., Cristià, A., Seidl, A., & Johnson, E. K. (2010). Developmental Changes in Infants' Ability to Cope with Dialect Variation in Word Recognition. Infancy, 15(6), 650–662.

Schuler, K. D., Yang, C., & Newport, E. L. (2016). Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. In Papafragou, A., Grodner, D. J., Mirman, D., & Trueswell, J. (eds.), Proceedings of the 38th annual conference of the cognitive science society. Austin, TX: Cognitive Science Society.

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. Psychological Science, 12(4), 323–328.

Smith, J., Durham, M., & Richards, H. (2013). The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation. Linguistics, 51(2), 285–324.

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.

Sundara, M., Demuth, K., & Kuhl, P. K. (2011). Sentence-position effects on children's perception and production of English third person singular –s. Journal of Speech, Language, and Hearing Research.

Tagliamonte, S. A., & D'Arcy, A. (2007). The modals of obligation/necessity in Canadian perspective. English World-Wide, 28(1), 47-87.

Tagliamonte, S. A., & D'Arcy, A. (2009). Peaks Beyond Phonology: Adolescence, Incrementation, and Language Change. Language, 85(1), 58–108.

Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. Journal of Child Language, 28(1), 127-152.

Traugott, E. C. (1989). On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change. Language, 65(1), 31–55.

Traugott, E. C. (2006). Historical aspects of modality. In W. Frawley (Ed.), The Expression of Modality (pp. 107–139). Berlin: De Gruyter.

Traugott, E. C., & Dasher, R. B. (2002). Regularity in semantic change. Cambridge, UK: Cambridge University Press.

Ünal, E., & Papafragou, A. (2016). Production–comprehension asymmetries and the acquisition of evidential morphology. Journal of Memory and Language, 89, 179-199.

van Dooren, A., Dieuleveut, A., Cournane, A., & Hacquard, V. (2017). Learning what must and can must and can mean. In Proceedings of the 21st Amsterdam Colloquium. Amsterdam, The Netherlands, December 20-22, 2017.

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). Somerville, MA: Cascadilla Press.

van Gelderen, E. (2011). The Linguistic Cycle. Oxford, UK: Oxford University Press.

Visser, F. T. (1963). An historical syntax of the English language. Leiden: Brill Archive.

von Fintel, K., & Gillies, A. S. (2010). Must... stay... strong! Natural Language Semantics, 18(4), 351-383.

Wagner, L. (2001). Aspectual influences on early tense comprehension. Journal of child language, 28(3), 661-681.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), Directions for Historical Linguistics (pp. 95-195). Austin, TX: University of Texas Press.

Werner, T. (2006). Future and non-future modal sentences. Natural Language Semantics, 14(3), 235-255.

Yang, C. (2016). The price of linguistic productivity: How children learn to break the rules of language. Cambridge, MA: The MIT Press.

Yanovich, I. (2013). Four pieces for modality, context and usage (Doctoral Thesis). Massachusetts Institute of Technology, Cambridge, MA, USA.