

## **Acquiring the Force of Modals: Sig you Guess What *sig* Means?**

**Anouk Dieuleveut, Annemarie van Dooren, Ailís Cournane, and Valentine Hacquard**

### **1. Introduction**

This paper investigates how children learn the *force* of their modals: how they map {*can/may/might*} to possibility meanings and {*must/need/have to*} to necessity meanings.

- |                            |   |
|----------------------------|---|
| (1) a. You <b>can</b> run. | possibility: ‘you are allowed to run’ (weak)  |
| b. You <b>must</b> run     | necessity: ‘you are required to run’ (strong) |

When and how do children figure out the force of their modals? We present results from (i) a corpus study of English, assessing how modal force is distributed in productions of 2 year-old children and in their input and (ii) from a Human Simulation Paradigm (HSP) study, assessing the informativity of natural contexts regarding force. After briefly discussing the various factors that might make it challenging to learn force from the input, we focus on the interplay of modals and negation.

Most prior studies of children’s comprehension of modals focus on Scalar Implicatures. Research has shown that by 4 years-old, children are sensitive to the relative force of their modals (Hirst & Weil 1982, Ozturk & Papafragou 2015, Moscati, Zhan & Zhou 2017, a.o.), but they seem to have difficulties when the contrast is not salient. First, they tend to over-accept possibility modals when necessity modals would be more appropriate (Noveck 2001, Ozturk & Papafragou 2015, ao). Second, they also over-accept necessity modals in possibility situations (Noveck 2001, Ozturk & Papafragou 2015, Koring 2018). While the former result could be attributed to children not deriving Scalar Implicatures (the inference that if you *can*, you *don’t have to*), which might be due to difficulties accessing the right scales (Barner, Brooks & Bale 2011, Bale & Barner 2013), the latter result is more surprising: while possibility modals are logically true (albeit under-

---

\* Dieuleveut, van Dooren, Hacquard, University of Maryland; Cournane, New York University. Corresponding email: adieulev@umd.edu (Dieuleveut). We thank our research assistants Joon Lee, Jan Michalowski and David Whitcomb, and the UMD/NYU ModSquad & Language Acquisition Lab for relevant comments and discussion. This project is supported by NSF grant #BCS-1551628.



1998, Courname 2015a, van Dooren & al. 2017). Do children realize that the modals used in comprehension studies express epistemic possibility and necessity? After all, having figured out for instance that *must* expresses deontic necessity does not necessarily entail knowing it expresses epistemic necessity.

## 1.2. Interaction with negation

These three aspects of the semantics and the pragmatics of modals conspire to make learning the force of individual modals from the input potentially challenging. How do children eventually figure out force? Here, we explore the role that negation might play in the learning process.<sup>2</sup>

Under negation (and more generally in any Downward-Entailing environment), logical relationships are reversed: ‘not possible’ (4b) is logically stronger than ‘not necessary’ (4a).

- (4) a. You don’t have to run.                      not > necessary    (weak)  
       b. You can’t run.                                not > possible     (strong)

Negated possibility modals thus have ‘strong’ meanings (Horn 2012). If we assume that ‘impossibility’ situations (such as prohibitions, for deontic flavor) are particularly salient and easy to identify from natural contexts, negation on possibility modals could help the child identify possibility modals as such.

In the case of necessity modals however, problems arise because scopal interactions between necessity modals and negation are not well-behaved (Iatridou & Zeijlstra 2013). In particular, *must* and *should* scope over negation, and are thus truth conditionally equivalent to negated possibility modals: (5a) is logically equivalent to (5b). This makes negation a reliable cue for (root) possibility modals, but not for all necessity modals.<sup>3</sup>

- (5) a. You mustn’t/shouldn’t run.            necessary > not    (strong)  
       b. You can’t run.                             not > possible     (strong)

In the following, we present results from (i) a corpus study, assessing how possibility and necessity modals are naturalistically distributed in early productions of children and in their input, focusing on their interaction with

---

<sup>2</sup> A number of comprehension studies also focus on children’s interpretation of sentences containing a modal and negation, that are potentially ambiguous between a ‘weak’ (‘possible that not’, which is logically equivalent to ‘not necessary that’) and a ‘strong’ interpretation (‘not possible that’/ ‘necessary that not’) (e.g. Moscati & Crain 2014, Koring Meroni & Moscati 2018, a.o.). This is not our focus here.

<sup>3</sup> This further interacts with modal flavor: for example, epistemic *may* scopes over negation but root *may* scopes under negation:

- (i) Ann may not be home  
     a. epistemic: ‘it is possible that Ann is not home’    possible > not (weak)  
     b. root: ‘it is not possible that Ann is home’        not > possible (strong)

negation and (ii) a Human Simulation Paradigm (HSP) study, using contexts from the corpus, to assess the informativity of natural contexts regarding force, and test the role of negation. Corpus results show that possibility modals may be acquired early on: children use them productively, both with and without negation, which we take as potential evidence that they master their force. On the other hand, children produce few necessity modals, despite the fact that these are more frequent than possibility modals in the input, which might suggest they are mastered later on. This is especially true for epistemic necessity modals.

The goals of the HSP were to (a) investigate how clear natural contexts are about force, and (b) test the effect of negation, our expectation being that negation leads to more informative contexts. Results show that adults are overall quite good at guessing modal force from natural contexts, with no significant difference between necessity and possibility. However, we found no support for the hypothesis that negative contexts are more informative than positive ones: negation actually decreased accuracy, though as we discuss, this may be due to an experimental artifact. Thus, a negative context may not be very informative on its own, but only in contrast with positive ones.

## 2. Corpus study

### 2.1. Goals and methods

The goal of this study was twofold: (a) to investigate children's early productions of modals focusing on the *force* dimension (most existing corpus studies of early production of modals focus on the *flavor* dimension, e.g. Papafragou 1998, Cournane 2015a, van Dooren & al. 2017; with the exception of Jeretic 2018 for French and Spanish) and (b) to explore the input children hear: how clear and informative is it for force?

We used the Manchester Corpus (Theakston et al., 2001) of UK English (CHILDES, MacWhinney 2000). This corpus consists of 12 child-mother pairs (6 females; age range: 1;09-3;00) recorded in unstructured play sessions. All modal utterances (45,151 of 564,625 total utterances; adult: 38,320; child: 6,831; excluding repetitions: adult: 37,194; child: 5,689) were coded for force (*possibility* vs. *necessity*) (6), negation (7) and flavor (8).

- (6) Functional modal lemmas by force
  - a. Possibility: *can, could, might, may; able to*
  - b. Necessity: *will, shall, would, must, should, need; have to, got to, supposed to, need to*
- (7) Negation
  - a. no negation: 'I **can** go to the pub now.'
  - b. negation:
    - on the modal: 'I **can't** get it'/ 'I must **not** forget Whispy.'
    - on higher auxiliary: 'we **don't** have to play with your toys.'
    - on embedding verb: 'I **don't think** you have to look for it.'
    - other negative quantifier: '**nobody** can reach it.'

- (8) Flavor
- a. epistemic: ‘it **must** be morning time.’
  - b. root: ‘you **mustn**’t eat things until you’ve paid for them.’
  - c. future: ‘well it **will** hurt.’

Note that we don’t distinguish ‘weak’ necessity modals (e.g. *should*) from necessity modals. We include *able* as a (possibility) modal. We considered *will* as a (necessity) modal, even if its modal nature is debated (Kissine 2008, Klecha 2016, a.o.), and decided to create a third ‘future’ category besides the root and epistemic category for flavor for *will*, *would* and *shall*. Note that we don’t differentiate amongst various root flavors (e.g. ability, teleological, deontic).

## 2.3. Results

### 2.3.1. Modal lemma by force

We find a strong asymmetry between children’s and adults’ production of possibility and necessity modals: while necessity modals are more frequent than possibility in the input, children produce more possibility than necessity modals (child: 32.4% necessity modals; adult: 60.6%) (Table 1).

**Table 1. Counts and percentages of modal uses by force, ordered by lemma frequency, for adults and children (repetitions excluded: 5.0% of the data)<sup>4</sup> (glmer: Force~AgeGroup+(1|CorpusPair):  $\beta = 1.16, p < .001^{***}$ )**

	adult (n=36847)	child (n=5686)
<b>possibility (P)</b>	<b>14500 (39.3%)</b>	<b>3841 (67.5%)</b>
<i>can</i>	11478 (31.1%)	3663 (64.4%)
<i>could</i>	1451 (3.9%)	86 (1.5%)
<i>might</i>	1217 (3.3%)	80 (1.4%)
<i>able</i>	315 (0.8%)	3 (0.0%)
<i>may</i>	39 (0.1%)	9 (0.1%)
<b>necessity (N)</b>	<b>22347 (60.6%)</b>	<b>1845 (32.4%)</b>
<i>will</i>	9911 (26.8%)	481 (8.4%)
<i>shall</i>	4739 (12.8%)	319 (5.6%)
<i>have to</i>	2399 (6.5%)	352 (6.1%)
<i>would</i>	2199 (5.9%)	42 (0.7%)
<i>got to</i>	941 (2.5%)	288 (5.0%)
<i>should</i>	793 (2.1%)	22 (0.3%)
<i>need (to)</i>	493 (1.3%)	217 (3.8%)
<i>must</i>	453 (1.2%)	114 (2.0%)
<i>supposed</i>	335 (0.9%)	9 (0.1%)
<i>ought to</i>	84 (0.2%)	1 (0.0%)

<sup>4</sup> We considered as repetitions cases where the speaker (child or adult) repeated a sentence uttered right before by herself or by another speaker.

### 2.3.2. Interaction with negation

We find that negation cooccurs more frequently with possibility than with necessity modals in the input (21.0% possibility with negation vs. 12.9% necessity with negation) (Table 2). This difference is significantly higher for children: 51.0% of possibility modals occur with negation, whereas 11.9% of necessity modals occur with negation. We find overall few modals with negative quantifiers (e.g. *nothing/never*), with no difference between possibility and necessity modals, as well as under negated embedding verb (e.g. *don't think*), again with no difference between possibility and necessity modals (possibility: 1.5%; necessity: 2.1%). Further details with a breakdown per modal is provided at [http://ling.umd.edu/~hacquard/project\\_modality.html](http://ling.umd.edu/~hacquard/project_modality.html).

**Table 2. Counts and percentages of modal uses, by force, with and without negation, for adults and children (excluding tag-questions<sup>5</sup> and repetitions) (glmer: Force~Negation+(1|CorpusPair): adult:  $\beta = 0.602, p < .001^{***}$ ; child:  $\beta = 2.03, p < .001^{***}$ )**

	adult (n=34191)		child (n=5605)	
	no negation	negation	no negation	negation
<b>P</b>	10683 (78.9%)	2850 (21.0%)	1861 (48.9%)	1938 (51.0%)
<b>N</b>	17993 (87.0%)	2665 (12.9%)	1591 (88.0%)	215 (11.9%)

Children's first productions of possibility modals (with and without negation) are found slightly earlier than necessity (Table 3).

**Table 3. Children's (n=12) first uses of possibility and necessity modals with and without negation, plotted within 2 month intervals.**

	Age range (n=12)				
	[2:00:2:02]	[2:02:2:04]	[2:04:2:06]	[2:06:2:08]	[2:08:2:10]
<b>P</b> no negation	10	1	1	-	-
negation	9	-	2	-	1
<b>N</b> no negation	6	2	2	2	-
negation	7	3	1	-	1

### 2.3.3. Interaction with flavor

We find that possibility and necessity modals are evenly distributed across flavor. (Following our coding schema, future is only expressed by necessity

<sup>5</sup> Tag questions (e.g., 'you can wash it later, **can't** you?') were very frequent in this corpus (adult: 7.1% of all modal utterances; possibility: 6.6%; necessity: 7.5%). We decided to exclude them, as we find an equal proportions across necessity and possibility modals.

modals). Epistemic uses are overall very rare (adult: 3.5%; children: 1.4%), particularly necessity epistemics (adult: 0.9% of all utterances; child: 0.3%).

**Table 4. Counts and percentages of modal uses, by force and flavor, for adults and children (excluding tag questions and repetitions) (glmer: Force~Flavor+(1|CorpusPair), epistemic vs. root (future excluded): adult:  $\beta = 0.02$ ,  $p = 0.76$  (NS); child:  $\beta = -0.33$ ,  $p = 0.20$  (NS))**

	adult (n=34183)			child (n=5601)		
	root	epistemic	future	root	epistemic	future
P	12665 (37.0%)	868 (2.5%)	na	3740 (66.7%)	59 (1.0%)	na
N	5014 (14.6%)	341 (0.9%)	15301 (44.7%)	981 (17.5%)	21 (0.3%)	804 (14.3%)

## 2.4. Corpus study Summary

We find a strong asymmetry between children's and adults' productions of modals: while adults produce more frequently necessity modals, children produce overall more possibility modals. Negation cooccurs more frequently with possibility than necessity modals in adult's speech, a difference which is even stronger for children who use possibility modals *with* negation (e.g. *can't*) more often than without (*can*). Last, while possibility and necessity modals show similar distribution for root and epistemics flavors, epistemics are much less frequent overall: necessity epistemic modals are thus particularly rare.

## 3. Human Simulation Paradigm (HSP)

### 3.1. Method and goals

In order to test (a) the informativity of natural contexts regarding force and (b) the potential effect of negation, we ran a Human Simulation Paradigm study (HSP) (Gillette & al. 1999), using contexts from the Manchester corpus.

In the experiment, participants were asked to guess a missing modal in a dialogue between a child and her mother. All trials consisted of the modal sentence with a blank and the 7 preceding utterances (see Figure 1). The whole list of contexts used is provided online.

MOTHER: put the others in as well.  
 CHILD: put the others in.  
 MOTHER: put the lid on the box now.  
 CHILD: yeah.  
 CHILD: no more.  
 MOTHER: no.  
 MOTHER: there's only four I'm afraid.  
 MOTHER: the other two \_\_\_\_\_ have been used.

### Figure 1: Human Simulation Paradigm stimuli: example trial (*must*)

We tested 4 modals (POSSIBILITY: *can, could*; NECESSITY: *must, should*)<sup>6</sup> (between subjects) in 4 conditions of sentence-type (affirmative, negative, interrogative, tag-question) (within subjects).<sup>7</sup>

(9) Sentence type condition

- affirmative (baseline): 'you \_\_\_ be careful.'
- negative: 'you \_\_\_ n't bang it into your car.'
- interrogative: '\_\_\_ you remember?'
- tag-question: 'you \_\_\_ say you're sorry, \_\_\_ n't you?'

Overall, there were 4 trials per participant. Participants had to make a guess after each dialogue. They were told in the instructions that the word would be the same across the 4 dialogues. The 4 dialogues were randomly selected from the list of contexts extracted from the corpus (details on the extraction procedure below). Overall, participants had to guess 2 words: they first had a training on the definite article *the* (they were told in the instructions that the 2 words were not related). After the 4 trials, they were asked to give a confidence rating on their answer.

Extraction procedure: 620 contexts were randomly extracted from the corpus for the 4 modals (*can, could, must, should*) according to the 4 conditions of sentence type (affirmative, negative, interrogative, tag).<sup>8</sup> After extraction, they were manually checked. The only selection criterion was that we excluded the contexts where the child herself used the modal in the dialogue. In cases where the adult was using the same modal several times in the dialogue, we put several blanks; in cases where another modal was used in the dialogue (either by the child or the adult), we left it as such, regardless of its force.

## 3.2. Results

### 3.2.1. Participants

---

<sup>6</sup> We chose auxiliaries to avoid distributional difference between possibility and necessity modals. Specifically, we had to avoid semi-auxiliaries as *have\_to* as it is impossible to answer with a possibility modal given the frame 'you don't \_\_\_\_\_ be so nice'.

<sup>7</sup> We added 'interrogative' and 'tag-question' conditions in order to get rid of a potential confound, negation being very rare in interrogative and frequent in tag-questions. (Overall, 21.0% of adults declarative sentences are negated (P: 26.9%; N: 17.3%); but only 5.6% of interrogative (P: 10.0%; N: 3.0%) and 55.9% of tag-questions (P: 62.9%; N: 52.0%). We excluded negated interrogatives (e.g. 'can't we play a nicer game?') from the HSP.

<sup>8</sup> Due to the low frequency of *must* in interrogatives (n = 13 in the whole corpus) and in tag questions (n = 40), we had to use *should* only in these conditions. We adapted the number of participants accordingly. Details of the number of contexts is available online.

We recruited 570 participants (324 females; language: US English) via Amazon MTurk. They were randomly assigned to one of the 4 groups (POSSIBILITY: *can, could*; NECESSITY: *must, should*).

### 3.3.2. Informativity of context

Answers were considered correct if participants gave a necessity modal (N-answer) for a necessity modal, or a possibility modal (P-answer) for a possibility modal (e.g. answering *could* in a *can* context was considered correct).<sup>9</sup> Clear typos (e.g. *couuld* for *could*) were considered as correct. We assume, following previous HSPs (Gillette & al. 1999, White 2016), that accuracy reflects informativity of the context. Table 5 reports the mean proportion of possibility vs. necessity answers depending on force (accuracy thus corresponds to bolded cells). Table 6 gives the results restricting to the first trial only.

**Table 5. Type of answer given by the participant (P, N, X) by force (n=570)**

		P-answer	N-answer	X-answer
POSSIBILITY	<i>can</i>	<b>63.8%</b>	14.3%	21.9%
	<i>could</i>	<b>63.3%</b>	23.4%	13.3%
	total	<b>63.6%</b>	18.7%	17.8%
NECESSITY	<i>must</i>	14.9%	<b>69.1%</b>	16.1%
	<i>should</i>	15.5%	<b>68.4%</b>	16.1%
	total	15.3%	<b>68.6%</b>	16.1%

**Table 6. Type of answer given by the participant (P, N, X) by force (n=570), on first trial only**

		P-answer	N-answer	X-answer
POSSIBILITY	<i>can</i>	<b>52.8%</b>	22.9%	24.3%
	<i>could</i>	<b>50.3%</b>	18.4%	31.3%
	total	<b>55.5%</b>	27.7%	16.8%
NECESSITY	<i>must</i>	14.3%	<b>59.8%</b>	25.9%
	<i>should</i>	13.9%	<b>63.4%</b>	22.8%
	total	14.6%	<b>57.8%</b>	27.6%

As expected, we found that performance improved over the course of the 4 dialogues (linear mixed model to predict Answer (*correct* vs. *incorrect*), using Order as fixed variable and Subject and Item as random factors (following Barr 2013): 1<sup>st</sup> vs. 2<sup>nd</sup> trial:  $\chi^2(1) = 14.2, p = 0.00016$  \*\*\*; 2<sup>nd</sup> vs. 3<sup>rd</sup> trial:  $\chi^2(1) = 5.41, p = 0.019$  \*; 3<sup>rd</sup> vs. 4<sup>th</sup> trial:  $\chi^2(1) = 6.78, p = 0.0092$  \*\*). We found no

<sup>9</sup> Answers that were considered as possibility modals (P-answers): *can, could, might, may*; considered as necessity modals (N-answers): *should, must, will, would, shall, ought, gonna, gotta, need*; considered as 'X' (X-answers) : *have, be, do*; any other word or non-word.

significant difference in accuracy between the different lexical instantiations of the force condition (*can* vs. *could*:  $\chi^2(1) = 0.001, p = 0.97$ ; *must* vs. *should*:  $\chi^2(1) = 0.0048, p = 0.94$ ). We therefore report merged results in the following.

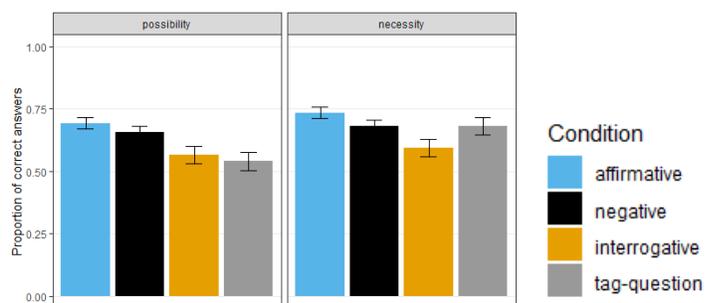
We tested force of guesses for each force condition, in order to see whether contexts were informative with respect to force, and found strong confirmation (P-answers in possibility vs. necessity contexts:  $\chi^2(1) = 270.3, p < .001^{***}$ ; N-answers:  $\chi^2(1) = 277, p < .001^{***}$ ). We then compared general accuracy in possibility vs. necessity contexts: participants were not significantly better at guessing for possibility than for necessity (linear mixed model with Subject and Item as random factors:  $\chi^2(1) = 3.17, p = 0.075$ ).

### 3.3.3. Effect of sentence type

Results by sentence-type condition are given in Table 7 (see Figure 2). We found that all non-declarative contexts decreased accuracy (linear mixed models testing effect of Condition on Accuracy, with Subject and Item as random factors: interrogative vs. affirmative:  $\chi^2(1)=16.083, p < .001^{***}$ ; negation vs. affirmative:  $\chi^2(1) = 4.3487, p=.037^*$ ; tag-question vs. affirmative:  $\chi^2(1)=12.2, p < 0.001^{***}$ ).

**Table 7. Accuracy by force and sentence type**

	affirmative	negative	interrogative	tag-question
POSSIBILITY	69.4%	65.8%	56.7%	54.0%
NECESSITY	73.6%	68.3%	59.5%	68.2%



**Figure 2. Accuracy by force and sentence type (n=570)**

### 3.3. Discussion

Overall, participants were quite good at guessing the force of the modal, suggesting that natural contexts are informative about force (at least for adults, who already have a modal inventory). We found no significant difference in accuracy between possibility and necessity contexts, but if any we find a trend in the direction of necessity contexts being easier (68.6% vs. 63.6%). Below, we discuss the effect of negation and interrogative sentences, that were both found to decrease accuracy.

### 3.3.1. Effect of negation

A priori, the result that negative contexts decrease accuracy seems to go against our expectations: that impossibility contexts would be easier to identify than possibility contexts, because potentially more salient (for example, a sentence containing a negated root possibility modal such as ‘you can’t touch the marbles.’ seems very easy to identify as a prohibition).

Before going further, note that our expectations differed for necessity modals and possibility modals: while for possibility modals (*can/could*), we expected negation to increase accuracy, for necessity modals (*must/should*), we had no clear expectation. This is because we had to test two modals that scope over negation (see footnote 8), and thus lead to impossibility interpretations as well as negated possibility modals (all negative contexts used in the experiment thus corresponded to impossibility situations, regardless of their force). Given that there is no way to compare directly necessity and impossibility in term of logical strength, we had no prior expectations on one being more informative than the other.

Putting our HSP and corpus results together, it appears that seeing a modal in a negative context might not be very informative on its own, but it may be useful to see the same modal frequently in both negative and affirmative contexts, as possibility modals seem to do in naturalistic data. Our current HSP paradigm didn’t allow us to test this possibility further. Note further that the decreased accuracy found in negative contexts may also come from an experimental artefact: in negative (but not affirmative) sentences, modal auxiliaries compete with *do*. Post-hoc analysis indeed revealed that most errors in negative contexts came from participants answering *do* (e.g. *do* for *must* in ‘\_\_\_n’t put crayons in your mouth’) (Table 8), thus not necessarily reflecting a real difference in informativity.

**Table 8 - Details of answers by force and condition (post-hoc analysis)**

		P-answer	N-answer	<i>do</i>	<i>be</i>	X (other)
P	affirmative	69.4%	20.5%	<b>5.6%</b>	0.3%	4.3%
	negative	65.8%	13.6%	<b>17.5%</b>	0.0%	3.1%
	interrogative	56.7%	19.1%	<b>20.6%</b>	1.5%	2.1%
	tag-question	54.0%	24.1%	<b>17.1%</b>	2.1%	2.7%
N	affirmative	15.8%	73.6%	<b>4.1%</b>	0.8%	5.7%
	negative	11.3%	68.3%	<b>13.6%</b>	3.0%	3.8%
	interrogative	21.1%	59.5%	<b>10.0%</b>	2.6%	6.8%
	tag-question	17.1%	68.2%	<b>9.4%</b>	3.5%	1.8%

### 3.3.2. Effect of interrogative and tag-questions

We also found that accuracy decreases in interrogative contexts. Here again, Table 8 shows that a potential explanation of this effect is the competition with *do* (e.g. answering *do* in ‘\_\_\_ you remember?’): results would then not show a real difference in informativity. We leave open the question of why (if the result

is not solely driven by an experimental artefact) interrogative contexts might be less clear about force, and for the higher accuracy found for necessity tag question.

Note that this paradigm is intrinsically limited, mainly because the task for the adult (guessing a word she already knows) is different from the task of a learner (guessing the meaning of a new word). We assumed, following previous HSP, that accuracy reflects informativity of contexts, but we did not control for example for the effect of ‘idioms’ or somewhat conventionalized expressions (e.g. ‘\_\_\_ I have another one please?’).

#### 4. General Discussion

*When* do children figure out the force of their modals? The primary goal of our corpus study was to investigate the early productions of children and their input, focusing on the force dimension (as prior corpus studies most focus on the flavor dimension, and comprehension studies usually test children older than 4 years old). We found that children produce possibility modals frequently and productively, both with and without negation, which we take as potential evidence that they master their force early on. Note that the same pattern seems to be found in other languages as well (for French and Spanish, see Jeretic (2018), and for Dutch, see van Dooren & al. (in progress)). Could negation play a facilitating role? Our HSP results suggest that a negative context, on its own, may not be more informative than an affirmative one. But it may be useful to see the same modal both with and without negation, which occurs frequently with possibility modals, but not necessity ones in the input.

Children’s low production of necessity modals *cannot* be explained by a lack of input, given that our corpus results show that adults actually produce more necessity modals. This result may in part be explained by differences in parents’ and children’s preferred topics of conversation, for example parents talking more about necessities (e.g. giving more orders). However, it could also suggest that children haven’t yet mastered necessity modals. Such a delay could explain the difficulties previously described in comprehension studies (Noveck 2001, Ozturk & Papafragou 2015): if children haven’t fully grasped necessity modals, they will both accept possibility modals in situations where necessity modals are more appropriate - since they lack the appropriate scale-mate, *and* necessity modals in possibility contexts - since they have not yet grasped their underlying force. This seems particularly relevant for epistemic modals, which are most often targeted in the experimental literature: epistemic necessity modals are very rare in input (0.9%), and thus likely to be mastered later. If children don’t know epistemic necessity modals, they are unlikely to compute the relevant scalar implicature.

Our HSP results suggest that children’s apparent delay with necessity modals does not stem from necessity meanings being harder to identify as such in natural contexts: we find no significant difference in accuracy between necessity and possibility contexts, but if any, it is in the direction of *necessity* being easier to guess.

How do children eventually figure out the force of necessity modals? This question will be addressed in future research. One possibility is that children use their early mastery of words that express possibility, to reason that necessity modals express some other force via mutual exclusivity. But this learning strategy might be limited by three facts: (i) in the modal domain, the same meaning is expressed by different words (*must*, *have to*), and the same words can express different meanings (e.g., epistemic and deontic necessity for *must*); (ii) because the same modals express different flavors in English, figuring out that a word like *can* expresses ability doesn't automatically guarantee that that word can express deontic possibility, which they could use to figure out the meaning of deontic *must*, for instance; (iii) Children may not necessarily expect that modals come in duals (e.g., Nez Perce only has a possibility modal, see Deal 2011).

## References

- Bale, Alan, and David Barner. "Grammatical alternatives and pragmatic development." *Alternatives in semantics*. Palgrave Macmillan, London, 2013. 238-266.
- Barner, David, Neon Brooks, and Alan Bale. "Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference." *Cognition* 118.1 (2011): 84-93.
- Barr, Dale J. "Random effects structure for testing interactions in linear mixed-effects models." *Frontiers in psychology* 4 (2013): 328.
- Courmane, Ailís. "Revisiting the epistemic gap: Evidence for a grammatical source." *Proceedings of the 39th annual BUCLD*. 2015.
- Deal, Amy Rose. "Modals without scales." *Language* (2011): 559-585.
- Gillette, Jane, Henry Gleitman, Lila Gleitman, and Anne Lederer. "Human simulations of vocabulary learning." *Cognition* 73, no. 2 (1999): 135-176.
- Gleitman, Lila R., and Barbara Landau, eds. *The acquisition of the lexicon*. MIT Press, 1994.
- Gleitman, Lila R., et al. "Hard words." *Language Learning and Development* 1.1 (2005): 23-64.
- Grice, H. Paul. "Logic and conversation." *1975* (1975): 41-58.
- Hirst, William, and Joyce Weil. "Acquisition of epistemic and deontic meaning of modals." *Journal of child language* 9.3 (1982): 659-666.
- Horn, Laurence. "Histoire d'\*O: Lexical pragmatics and the geometry of opposition." *New perspectives on the square of opposition* (2012): 393-426.
- Horn, Laurence. "On the semantic properties of logical operators in English: UCLA dissertation." *Distributed by IULC*(1972).
- Iatridou, Sabine, and Hedde Zeijlstra. "On the scopal interaction of negation and deontic modals." *Logic, language and meaning*. Springer, Berlin, Heidelberg, 2010. 315-324.
- Jeretič, Paloma. 2018. Evidence for children's dispreference for root modal weakness. Ms. New York University.
- Kissine, Mikhail. "Why will is not a modal." *Natural Language Semantics* 16.2 (2008): 129-155.
- Klecha, Peter. "Diagnosing modality in predictive expressions." *Journal of Semantics* 31.3 (2013): 443-455.
- Koring, Loes, Luisa Meroni, and Vincenzo Moscati. "Strong and Weak Readings in the Domain of Worlds: A Negative Polar Modal and Children's Scope Assignment." *Journal of psycholinguistic research* (2018): 1-25.

- Kratzer, Angelika. "What 'must' and 'can' must and can mean." *Linguistics and philosophy* 1, no. 3 (1977): 337-355.
- Kuczaj, Stan A., Michael P. Maratsos. "What children can say before they will." *Merrill-Palmer Quarterly of Behavior and Development* 21, no. 2 (1975): 89-111.
- MacWhinney, Brian. The CHILDES project: Tools for analyzing talk, Volume I: Transcription format and programs. Psychology Press, 2014.
- Moscati, Vincenzo, and Stephen Crain. "When negation and epistemic modality combine: the role of information strength in child language." *Language Learning and Development* 10.4 (2014): 345-380.
- Moscati, Vincenzo, Likan Zhan, and Peng Zhou. "Children's on-line processing of epistemic modals." *Journal of child language* 44.5 (2017): 1025-1040.
- Noveck, Ira A. "When children are more logical than adults: Experimental investigations of scalar implicature." *Cognition* 78.2 (2001): 165-188.
- Ozturk, Ozge, and Anna Papafragou. "The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation." *Language Learning and Development* 11.3 (2015): 191-214.
- Papafragou, Anna. "The acquisition of modality: Implications for theories of semantic representation." *Mind & language* 13.3 (1998): 370-399.
- 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/>.
- R Core Team (2013). R: A language and environment for statistical computing. R
- Searle, John R. "Indirect speech acts." (1975): 59-82.
- Theakston, Anna L., Elena VM Lieven, Julian M. Pine, and Caroline F. Rowland. "The role of performance limitations in the acquisition of verb-argument structure: An alternative account." *Journal of child language* 28, no. 1 (2001): 127-152.
- White, Aaron Steven, Valentine Hacquard, Philip Resnik, and Jeffrey Lidz. "The contextual modulation of semantic information."