Devoir, ou *pouvoir,* that is the question

Modals can express different forces: possibility (e.g., "you *can"*) or necessity (e.g., "you *must"*). Modal force raises a *Subset problem* for learners [1,2,3]: given that necessity entails possibility, how can children realize that *must* is stronger than *can*? Existing acquisition studies suggest that children struggle with necessity modals particularly, contrasting with their early mastery of possibility modals [3,4,5]. Yet, most of these studies focus on English, where necessity modals are much rarer than possibility modals in talk to children (children's "input") [3], suggesting that the delay could just be due to a lack of exposure. In this study, we show by looking at French, that this isn't sufficient: despite more exposure, French children also struggle with necessity modals.

Background. [3] ran a corpus study of children's modal productions and input from the Manchester Corpus [6] (CHILDES database [7]). They show that English children use possibility modals like *can* early, frequently, but struggle with necessity modals like *must*/*have to*, using them later on, less frequently, and crucially not in an adult-like way. To make results directly comparable, we stayed as close to [3]'s methods as possible, applying them to a French corpus.

Corpus. We used the Lyon Corpus [8] (5 children; age range: 1;00-3;00), and the Paris corpus [9] (6 children; age range: 0;7-6;03), on CHILDES [7]. We extracted and coded modal utterances for force (Possibility: *pouvoir*; Necessity: *devoir*/*falloir*/*avoir-à*), type of modality ('epistemic' vs 'root'), and negation. **Results**. Modal utterances represent 3.8% of all adult utterances (vs 5.8% in English), and 1.9% of child utterances between age 2 and 3. **Table 1** summarizes counts of adult and child productions comparing French and English. We find that in French adult talk, necessity modals are more frequent (62% of all their modal utterances, vs 28% in English). Yet, French children produce more possibility (62%). As in English, they also produce possibility modals earlier (mean age of 1st production: *pouvoir*=1:11; *falloir*=2:03; *devoir*=2:11; *avoir-à*=5:06).

Experiment 1. To test child usage, we use a paradigm introduced by [3]. Its goal is to determine whether children use necessity and possibility modals in an adult-like way. (Adult) participants are presented with mother-child dialogues extracted from the corpus and asked to guess the force of a blanked out modal, by picking between two options: either a possibility (*pouvoir*) or a necessity modal (*devoir*/*falloir*). The modal is uttered either by a child (**Fig1-i**) or by her mother (**Fig1-ii**). **Procedure**. All experiments were coded with PCIbex and run online. Overall, participants had to judge 40 dialogues, presented in a randomized order (20 controls, 20 trials: 10 possibility, 10 necessity, randomly selected out of a list of 20 dialogues randomly extracted from the corpus). **Conditions**. We had three groups based on the speaker's age: 2-3-year-olds, 4-5-year-olds, Adults (used as baseline). We ran two versions varying the necessity modal (**Exp1_*d***: *devoir*; **Exp1_*f***: *falloir;* we don't test *avoir à* because it is too rare). We test only 'root' modals because epistemic uses are too rare in children's production ([10]), and we excluded negated utterances to avoid issues from the scopal interaction of modals and negation ([11]). Force was tested within subject, Age and Lemma between subjects. **Participants.** 358 French participants were recruited on Prolific (60 per condition, 2 failed to record data) (166 F, 186 M, 6 NB; mean age: 32.8yrs). We removed 11 participants whose accuracy scores on controls was <75% (3.1%) (**Exp1_*d***: ADU: 59; CHI2-3: 56; CHI4-5: 59; **Exp1_*f***: ADU: 59; CHI2-3: 54; CHI4-5: 58). **Results. Fig2** summarizes the mean accuracy for each condition. We use generalized linear mixed effects models, built with a maximal random effect structure, testing Accuracy (dependent variable, binomial), with Force as fixed effect and Subject and Item as random factors, and compare them with reduced models without Force as a fixed effect [12,13]. ***Effect of Force.*** For adult production, participants are accurate at guessing force with no difference between possibility and necessity contexts (general mean accuracy: P: 78%; N: 77%). For child production, we find higher performance on possibility than necessity in both age groups (2-3yo: P: 75%; vs N: 60%; **Exp1_*d***: $\chi^2(1)=4$, $p=.04$*, **1_*f***: $\chi^2(1)=4.1$, $p=.04$; 4-5yo: P: 82% vs N: 64%; **1_*d***: $\chi^2(1)=5.3$, $p=.02$*; **1_*f***: $\chi^2(1)=6.5$, $p=.01$*). ***Effect of Age***. Comparing Child groups to Adult, we find significantly lower accuracy for necessity contexts in all age groups. For possibility contexts, we find a difference in **Exp1_*d***, but not in **1_*f***.

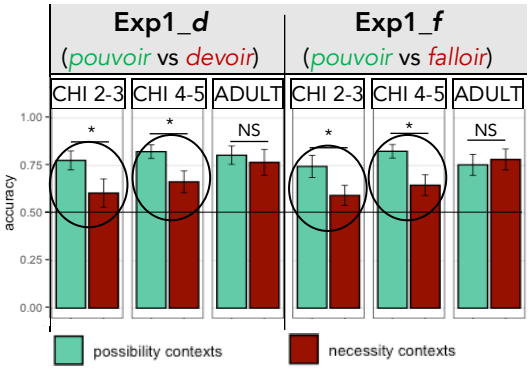
Figure 1. Experimental stimuli: example trials (*pouvoir* vs *devoir*)

(i) Exp1: Children's production	**(ii) Exp1**: Mothers' production	**(ii) Exp2**: With role reversal
ENFANT : ... t'en laisses un petit coup. MAMAN : merci. ENFANT : voilà. MAMAN : merci. ENFANT : arrête d'aller là avec le ptit chevaux ENFANT : vous arrêtez d'aller là. ENFANT : parce que c'est après. ENFANT : qu'on _____ aller après. peut doit	AUTRE ADULTE : oui. MAMAN : oui AUTRE ADULTE : une soucoupe. ENFANT : sont un peu vieilles. MAMAN : oui sont un peu abîmées tordues. MAMAN : ah celle-là elle marche bien. AUTRE ADULTE : merci beaucoup. MAMAN : tu _____ souffler dessus. peux dois	MAMAN : ... t'en laisses un petit coup. ENFANT : merci. MAMAN : voilà. ENFANT : merci. MAMAN : arrête d'aller là avec le ptit chevaux MAMAN : vous arrêtez d'aller là. MAMAN : parce que c'est après. MAMAN : qu'on _____ aller après. peut doit
CHILD: you leave a small bit/ MOTHER: thank you/ CHILD: Here you go/ MOTHER: Thank you/ CHILD: Stop going there with the little horse/ CHILD: You stop going there/CHILD: cause it's after/CHILD: that we ___ go after.	OTHER ADULT: Yes/ MOTHER: Yes/ ADULT: a plate/ CHILD: They're kind of old/ MOTHER: Yes they're a little TRANSLATE/ MOTHER: oh, this one works well/ ADULT: Thanks a lot/ MOTHER: you ____ blow on them.	MOTHER: you leave a small bit/ CHILD: thank you/ MOT: Here you go/ CHILD: Thank you/ MOT: Stop going there with the little horse/ MOT: You stop going there/ MOT: cause it's after/ MOTHER: that we ____ go after.

Table 1. Counts and percentage of modal uses in French and English, by force and speaker

		2-3-year-olds	3-5-year-olds	Adults
		count (%mod utt)	count (%mod utt)	count (%mod utt)
French	POSSIBILITY	850 (62%)	516 (58%)	2008 (38%)
	NECESSITY	529 (38%)	370 (42%)	3108 (62%)
	falloir	492 (36%)	298 (34%)	2659 (53%)
	devoir	21 (2%)	66 (7.4%)	403 (8%)
	avoir-à	16 (1%)	6 (1%)	46 (1%)
	ALL	**1379 (100%)**	**886 (100%)**	**5114 (100%)**
Eng	POSSIBILITY	3798 (79%)	Not assessed.	13500 (72%)
	NECESSITY	1002 (21%)		5353 (28%)
	ALL	**4800 (100%)**		**18853 (100%)**

Figure 2. Mean accuracy, Exp1 (n=347)



Experiment 2. We ran a follow-up study using the same dialogues, but switching the roles of child and mother, to see whether performance could come from some participants' expectations for children to use more possibility modals, rather than children effective misuses. **Fig1-iii** illustrates the manipulation. Half of the trials had the reversed speakers; half kept the original speaker, allowing us to replicate results from **Exp1**. We excluded contexts when role reversal was too odd, based on a naturalness rating with French naive participants (prop. excluded: 52%). We had four groups (**2_*d***: *pouvoir* vs *devoir*; **2_*f***: *pouvoir* vs *falloir*; judging either adult or child). From a participant's perspective, Exp2 was identical to Exp1. **Participants**. 120 French participants who hadn't taken part in Exp1 were recruited on Prolific (30 per condition) (66 M, 49 F, 2 NB, 3 unknown; mean age: 32.5yrs). We excluded 2 participants due to low accuracy on controls. **Results.** We find that adults' judgements remain stable: we replicate results from Exp1, both on unchanged dialogues (**Table 2**, row (**ii**) vs (**iii**)) and on role reversed contexts (row (**ii**) vs (**iv**)).

Table 2. Results (mean accuracy) of Exp2 (n=118)

		Exp_d (*pouvoir* vs *devoir*)				**Exp_f** (*pouvoir* vs *falloir*)			
		CHI (2-3yo)		ADULT		CHI (2-3yo)		ADULT	
		POSS	NECE	POSS	NECE	POSS	NECE	POSS	NECE
i	**Exp1** (all contexts)	78%	61%	80%	77%	75%	59%	75%	78%
ii	**Exp1** (kept in Exp2)	80%	66%	80%	72%	78%	57%	73%	79%
iii	**Exp2** (unchanged)	82%	65%	79%	72%	76%	59%	71%	85%
iv	**Exp2** (role reversed)	82%	66%	74%	64%	78%	54%	62%	83%

Discussion. We replicate [3]'s findings in French: children master possibility modals early, but struggle with necessity modals. They use them later on, less frequently, and crucially, don't use them in an adult-like way: they use them when adults expect possibility modals. While we still don't know the source of their difficulties, our study shows that they are not limited to English, and that "lack of exposure" can't explain them: French children actually hear more necessity than possibility modals in their input. Are children are confused about the meaning of necessity modals? Or, is it simply that they don't know yet in which contexts they are appropriate? These are questions to discuss, and call for extension to other logical scales where similar *Subset problems* arise, like *some/all* or *sometimes/always*.

References. [1] Berwick, 1985. [2] Wexler and Manzini, 1987. [3] Dieuleveut et al., 2022. [4] Noveck, 2001. [5] Ozturk and Papafragou, 2013. [6] Theakston, 2001. [7] MacWhinney, 2000. [8] Demuth and Tremblay, 2008. [9] Morgenstern and Parisse, 2007. [10] Cournane and Tailleur, 2020. [11]. Iatridou and Zeljstra, 2013. [12] Barr, 2013. [13] Team, R. 2013.