

Finding the force: a novel word learning experiment with modals

In English, modals express either possibility (e.g. *can*) or necessity (e.g. *must*). How do learners figure out that *must* is stronger than *can* (Horn scales [1])? And given that necessary entails possible, what prevents them from hypothesizing possibility meanings for necessity modals? Do learners exhibit a bias towards strong necessity meanings, to avoid a subset learnability problem [2,3]? We use a word learning experiment to test whether adults have force biases when learning novel modals, and whether these differ by ‘flavor’ context (epistemic (knowledge-based) versus teleological (goal-based)). As a control, we test actual English modals, using the same methods. We find no evidence for a necessity bias, but potentially a possibility bias in the teleological condition. Given significantly differing patterns for flavor contexts, we argue that the modal flavor dimension is an under-considered variable in modal force acquisition research.

Methods. Experiments were run using IBEX. Participants were introduced to Luke, who learned new words from a foreign dialect, Kabberton English, from native-speaker Mary. All participants learned 3 novel words in blocks. Word 1 was a control; Words 2 & 3 varied between subjects, either a new possibility modal first (learned in POSSIBILITY situations, tested in NECESSITY situations), or a new necessity modal first (learned in NECESSITY, tested in POSSIBILITY), followed by a block with the other modal. For all words, participants first saw Mary use the word (*training 1*: 4 items), then Luke (*training 2*: 4 *yes*-control, 4 *no*-control, with feedback), then the test phase (*test*: 6 trials, 6 *yes*-control, 6 *no*-control, no feedback). Participants had to judge whether Luke’s use of the word was correct (Yes/No). **Fig. 1** shows POSSIBILITY, NECESSITY and IMPOSSIBILITY situations by flavor, with test sentences in (1). We also ran the same experiment with English modals, testing *might/must* (epistemic) and *can/must* (teleological).

- (1) Sentences epistemic: ‘The keys **sig/gleeb** be in the [blue] box.’
 teleological: ‘We **sig/gleeb** go down the [blue] road.’

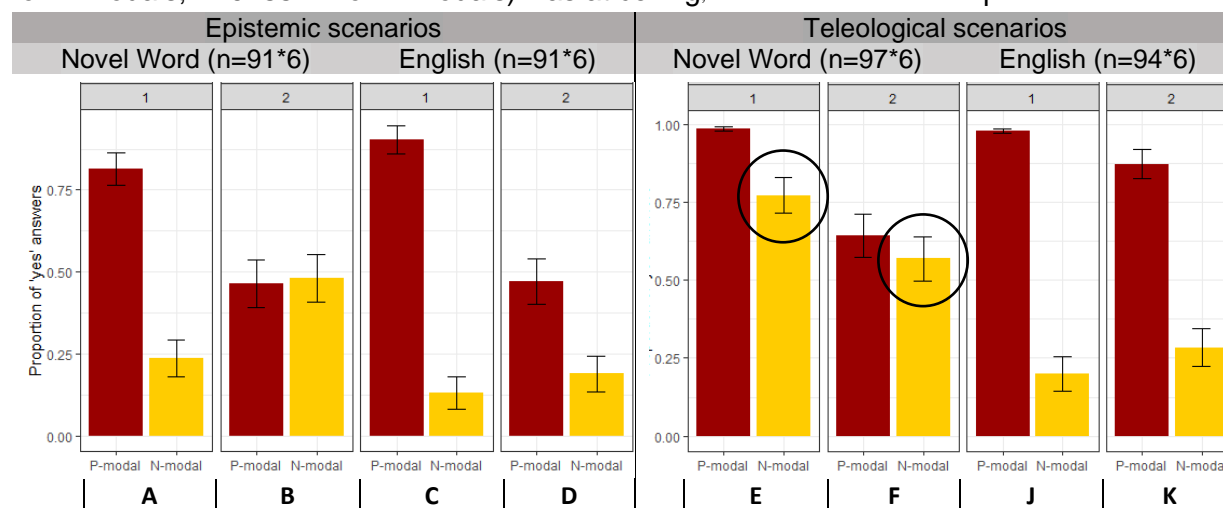
Figure 1. Visual stimuli and test sentence frames, for epistemic and teleological conditions

	POSSIBILITY situation	NECESSITY situation	IMPOSSIBILITY situation
epistemic The keys <u> </u> be in the <u>blue</u> box			
teleological We <u> </u> go down the <u>blue</u> road			

Conditions. Force (possibility vs. necessity) and Order (learned 1st vs. 2nd) were tested within-subjects, flavor (teleological vs. epistemic) between. **Participants.** 386 U.S. English participants were recruited on Amazon MT (Novel-Word (NW), n=194 (97 female, age *m*=37yrs); English (Eng): 192 (97 female, age *m*=38yrs); after exclusion (3.4%): 373 participants). **Expected results.** When learning a novel modal in POSSIBILITY situations, participants should assume a possibility meaning and accept it in NECESSITY, unless they know a stronger scalemate (Eng, or when tested 2nd in NW). When learning a modal in NECESSITY, participants could either assume possibility or necessity, and will correspondingly either accept or reject the modal in POSSIBILITY. **Results.** **Fig. 2** shows proportion of yes responses on test trials for possibility and necessity modals, depending on block order and flavor, for both NW and Eng experiments. **Analysis.** We use binomial linear mixed effects models, built with a maximal random effect structure, with Subject and Item as random factors [4,5]. In **epistemic** scenarios, we find that when learned 1st (see labels **A** and **C**), possibility modals are often accepted in NECESSITY (*sig*: 81.5%; *might*:

90.4%), with no significant difference between NW and Eng. Necessity modals are less often accepted in POSSIBILITY (*sig*: 23.6% vs. *must*: 13.0%), again with no difference between NW and Eng. In **teleological** scenarios (labels **E** and **J**), possibility modals learned 1st are also almost always accepted in NECESSITY (*sig*: 98.6%; *can*: 97.9%; NW vs. Eng.: $\chi^2(1)=0.46$, $p=.50$ (ns)). Necessity modals are accepted in POSSIBILITY less often, but there is a significant difference between NW and Eng. (*sig*: 77.2%; *must*: 19.9%; NW vs. Eng.: $\chi^2(1)=77.9$, $p<.0001$ ***). **Effect of Order** – When learned 2nd (labels **B**, **D**, **F** and **K**), we find the expected decrease for possibility modals in all conditions: they are accepted less often after learning/seeing a stronger scale-mate (epis: *sig*: 46.4%; *might*: 47.1%; tel: *sig*: 64.3%; *can*: 87.2%). The effect is significant for all conditions (epis: NW (**A vs. B**): $\chi^2(1)=10.1$, $p=.0015$ ***, Eng. (**C vs. D**): $\chi^2(1)=23.5$, $p<.0001$ ***; tel: NW (**E vs. F**): $\chi^2(1)=21.4$, $p<.0001$ ***; Eng. (**J vs. K**): $\chi^2(1)=8.1$, $p=.004$ **). With necessity modals, we find a slight increase for epistemics (**A vs. B** and **C vs. D**) (*sig*: 48.1%; *must*: 18.9%). For teleological (**E vs. J** and **F vs. K**), we find an increase for Eng, but a decrease for NW (*sig*: 56.9%; *must*: 28.4%). Order is significant for both Eng and NW (NW: $\chi^2(1)=36$, $p<.0001$ ***; Eng: $\chi^2(1)=38.9$, $p<.0001$ ***).

Figure 2. Proportion of yes responses for possibility (P) and necessity (N) modals in NECESSITY (red) and POSSIBILITY (yellow) situations, faceted for order (1st, 2nd) and flavor (Epistemic, Teleological) (n=373). Accuracy on controls (no controls: IMPOSSIBILITY; yes controls; POSSIBILITY for P-modals; NECESSITY for N-modals) was at ceiling, therefore we don't report it here.



Discussion. Participants accept novel possibility modals in NECESSITY situations at a very high rates, as expected when learned 1st (they don't know/consider a stronger alternative). They accept novel necessity modals in POSSIBILITY situations at a low rate for epistemics (23.6%), but at 77% for teleological scenarios, with a significant difference with English *must* (19.9%). Adults do not show a clear bias towards strong necessity meanings [2,3]: we would expect higher rejection rates. Thus, if child learners have such a bias, it does not seem to survive in adulthood. The high acceptance rate for novel necessity modals in teleological scenarios could be due to the saliency of an ability interpretation (the QUD is whether it is *possible or not*, instead of *possible or mandatory*), which would also explain the very high acceptance rate for English *can* in NECESSITY situations. Our results thus highlight that flavor variability matters when testing modal force. In particular, children's struggles with force reported in the acquisition literature [6,7] could arise in part from issues with flavor, rather than force or scalar implicatures.

References. [1] Horn, L. R. (1972). On the semantic properties of logical operators. [2] Berwick, R. C. (1985). *The acquisition of syntactic knowledge*. [3] Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. [4] Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. [5] Team, R. C. (2013). R: A language and environment for statistical computing. [6] Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. [7] Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality.