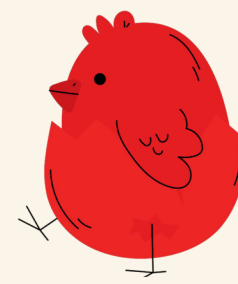
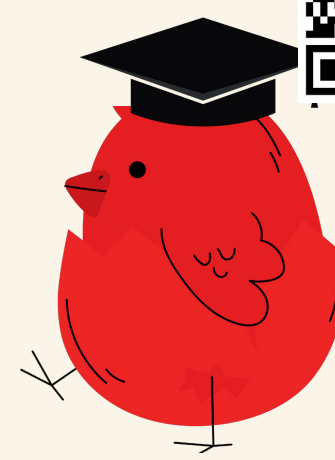


# Teacher Demonstrations in a BabyLM's Zone of Proximal Development

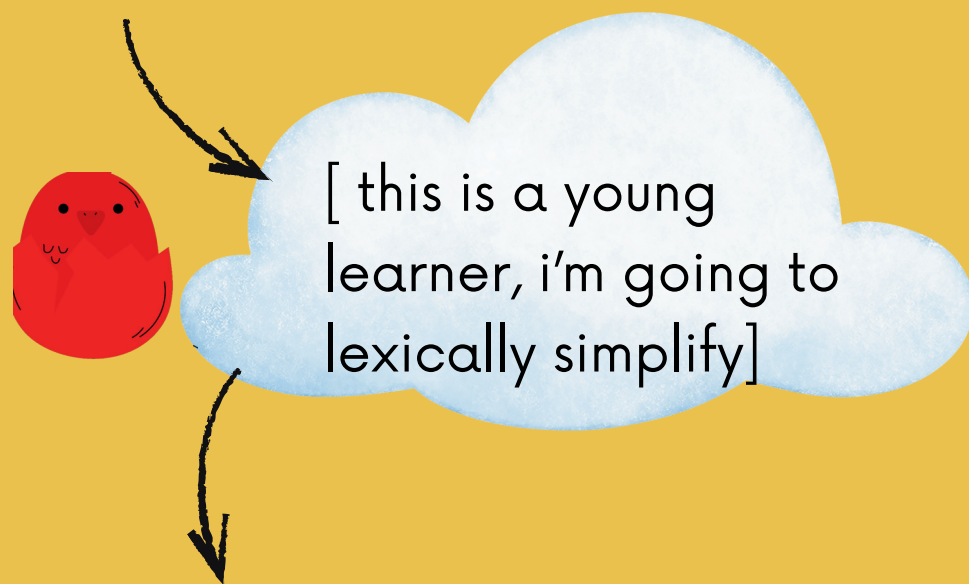
Suchir Salhan, Hongyi Gu, Donya Rooein, Diana Galvan-Sosa, Gabrielle Gaudeau, Andrew Caines, Zheng Yuan & Paula Buttery

arXiv



We develop **ContingentChat**, a Teacher-Student Post-Training Framework for BabyLMs to **improve multi-turn dialogic interaction**

## Contingency



The child is saying nonsense. I'll tell them to repeat

...

I give up! I'm going to tell the child the answer.

### Experiment 1: Post-Training

→ **CPO/ORPO + BabyLM**

**Llama 3.1B Teacher**

100M BabyLMs (OPT, seq len {1024, 4096})

iteratively guided via Teacher-edited preference pairs to produce higher-quality, contextually-appropriate responses.

## ContingentChat Alignment

30M words of BabyLM post-training data

Annotated Teacher-Student Dialogue

**5 dimensions** of discourse cohesion and complexity

### Experiment 2: Adaptive Decoding of Teacher Model

**BlenderBot 3B Teacher**

**BabyLM**

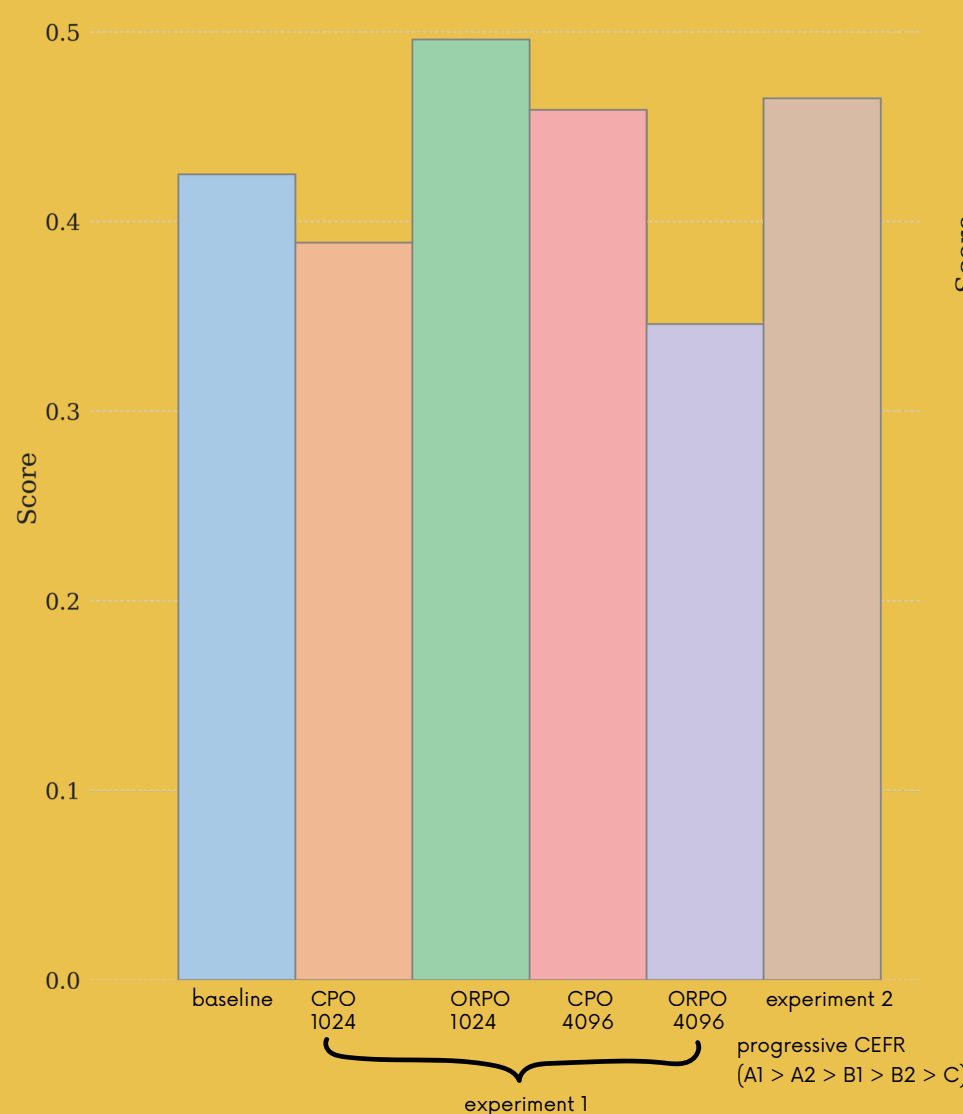
5 dialogues  
CEFR regressor



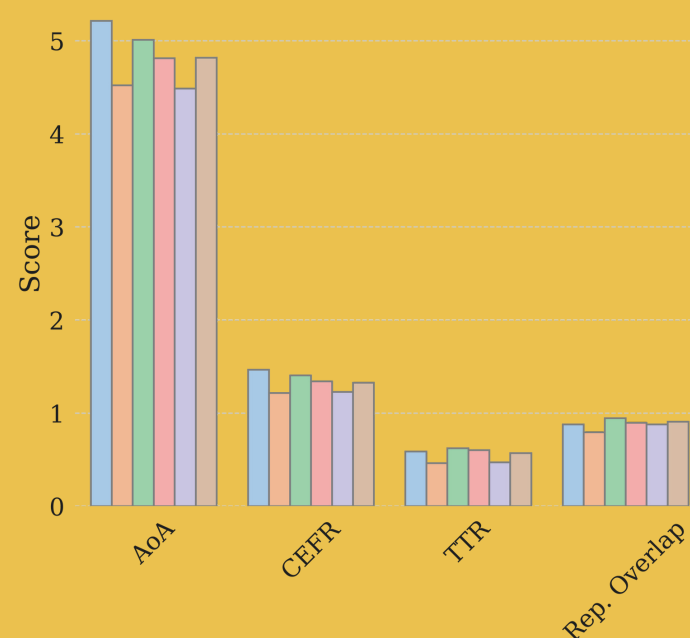
## ContingentChat Evaluation

### Coherence Metrics

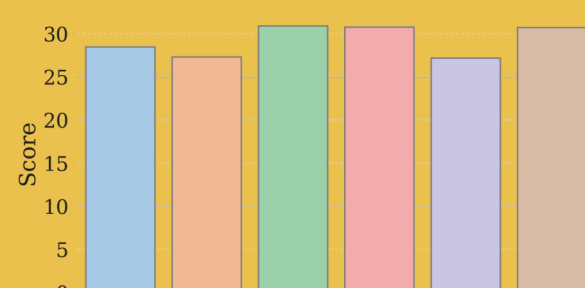
Average Cohesion Metric



Cohesion Evaluation Metrics



BabyLM Average Performance



opt-base cpo-opt-1024 orpo-opt-1024 opt-cefr-iteration5  
cpo-opt-4096

### Human Judgement

cpo-opt-1024 achieves the best overall performance

|               | GRM | WCH | COH | CNC | APP | COR |
|---------------|-----|-----|-----|-----|-----|-----|
| cpo-opt-1024  | ✓   | ✗   | ✓   | ✗   | ✗   | ✗   |
| cpo-opt-4096  | ✗   | ✗   | ✗   | ✗   | ✗   | ✗   |
| orpo-opt-1024 | ✓   | ✗   | ✓   | ✗   | ✗   | ✗   |
| orpo-opt-4096 | ✗   | ✗   | ✗   | ✗   | ✗   | ✗   |