Babysit A Language Model From Scratch: Interactive Language Learning by Trials and Demonstrations

** For Seminar Talk @ Deep Learning: Classics and Trends (DLCT)*

Martin Ziqiao Ma,*
Zekun Wang,*
Joyce Chai

COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF MICHIGAN

June 27th, 2024

Background

Social interaction is crucial in human language acquisition.

- The current dominant line of research focuses on *pragmatic inference*, i.e., children's ability to make inferences about people's communicative intents [1].
- There has been research looking into the role of *feedback* in interactive learning:
  - *Communicative Feedback* [2-3]: the explicit negotiation of shared understanding with the interlocutor to achieve and maintain common ground.
  - *Corrective Feedback* [4-5]: describing responses from caregivers that provide a correction for potential mistakes in children's utterances, and its variants such as negative evidence, reformulations, or recasts.

[1] Daniel Yurovsky, and Michael C. Frank. Beyond naïve cue combination: Salience and social cues in early word learning. Developmental Science 20.2 (2017): e12349.
[2] Mitja Nikolaus and Abdellah Fourtassi. Communicative feedback in language acquisition. New Ideas in Psychology. 68 (2023): 100985.
[3] Mitja Nikolaus, Laurent Prévot, and Abdellah Fourtassi. Communicative Feedback in Response to Children's Grammatical Errors. CogSci. 2023.
[4] Matthew Saxton. Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. First Language 20.60 (2000): 221-252.
[5] Matthew Saxton, Phillip Backley, and Clare Gallaway. Negative input for grammatical errors: Effects after a lag of 12 weeks. Journal of child language 32.3 (2005): 643-672.

Background

Corrective feedback in first language acquisition?

- Corrective feedback is shown to be helpful in second language acquisition [6], but…
- Researchers largely dispute its *availability* and *effectiveness* in human first language acquisition [7-8].
- In the context of AI models, corrective feedback can be massively available, and its effectiveness should be reconsidered. [Why?]
- Research Question:
  - Do corrective feedback benefit neural language (especially word) acquisition?
  - What are some phenomena under such training paradigm?

[6] El Tatawy, Mounira. "Corrective Feedback in Second Language Acquisition." Studies in Applied Linguistics and TESOL 2.2 (2002).
[7] Gary F Marcus. Negative evidence in language acquisition. Cognition 46.1 (1993): 53-85.
[8] E Mark Gold. Language identification in the limit. Information and control 10.5 (1967): 447-474.

Motivation

Connecting neural language models to language acquisition studies.

- Recently, several lines of cognitively motivated language modeling research have looked into the *learnability* and *learning efficiency* of language [9-10].
- Several efforts have explored potential mechanisms that contribute to efficient language learning in (vision-)language models by incorporating *non-linguistic inputs*, such as *multimodal stimuli* [11-12] and/or *communicative feedback* [13-14].

[9] Tyler A. Chang and Benjamin K. Bergen. Word acquisition in neural language models. Transactions of the Association for Computational Linguistics 10 (2022): 1-16.
[10] Evanson, Linnea, Yair Lakretz, and Jean Rémi King. "Language acquisition: do children and language models follow similar learning stages?." Findings of the Association for Computational Linguistics: ACL 2023. 2023.
[11] Shi, Haoyue, et al. "Visually Grounded Neural Syntax Acquisition." Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019.
[12] Ma, Ziqiao, Jiayi Pan, and Joyce Chai. "World-to-Words: Grounded Open Vocabulary Acquisition through Fast Mapping in Vision-Language Models." Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2023.
[13] Zhu, Hao, Yonatan Bisk, and Graham Neubig. "Simulated Language Learning from Communicative Goals and Linguistic Input." Proceedings of the Annual Meeting of the Cognitive Science Society. Vol. 44. 2022.
[14] Liu, Andy, et al. "Computational Language Acquisition with Theory of Mind." The Eleventh International Conference on Learning Representations. 2022.

Motivation

Connecting neural language models to language acquisition studies.

- How can neural language models be useful to language acquisition if they are not a priori designed to be cognitive models [15]? [More]
    - As proof-of-concept models: Models are independent learners from humans. They can show us what is possible 'in practice' for models and 'in principle' for humans.
    - As hypothesis generators: Models can also propose new hypotheses about language learning in children, which can themselves make testable predictions.
- What are the benefits?
    - Benefit 1: Easy *ablation studies* without ethical considerations.
    - Benefit 2: Easy access to model weights overtime, and can probe the model *at scale*.

[15] Eva Portelance. Neural Network Approaches to the Study of Word Learning. PhD Thesis, Stanford University.

Problem Formulation

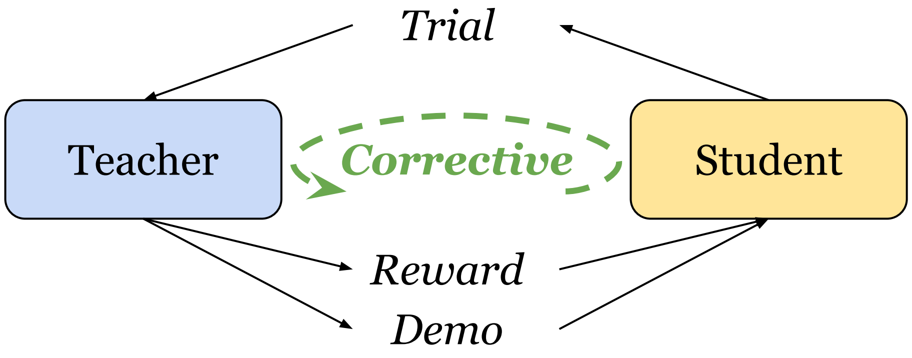
Corrective feedback in neural language models by trial and demo (TnD).

- Modeling corrective feedback in computational models presents challenges.
  - Recruiting human subjects to supervise the development of a language model from the ground up over numerous iterations is impractical;
  - The feedback takes the form of natural language rather than a simple heuristic score.

Problem Formulation

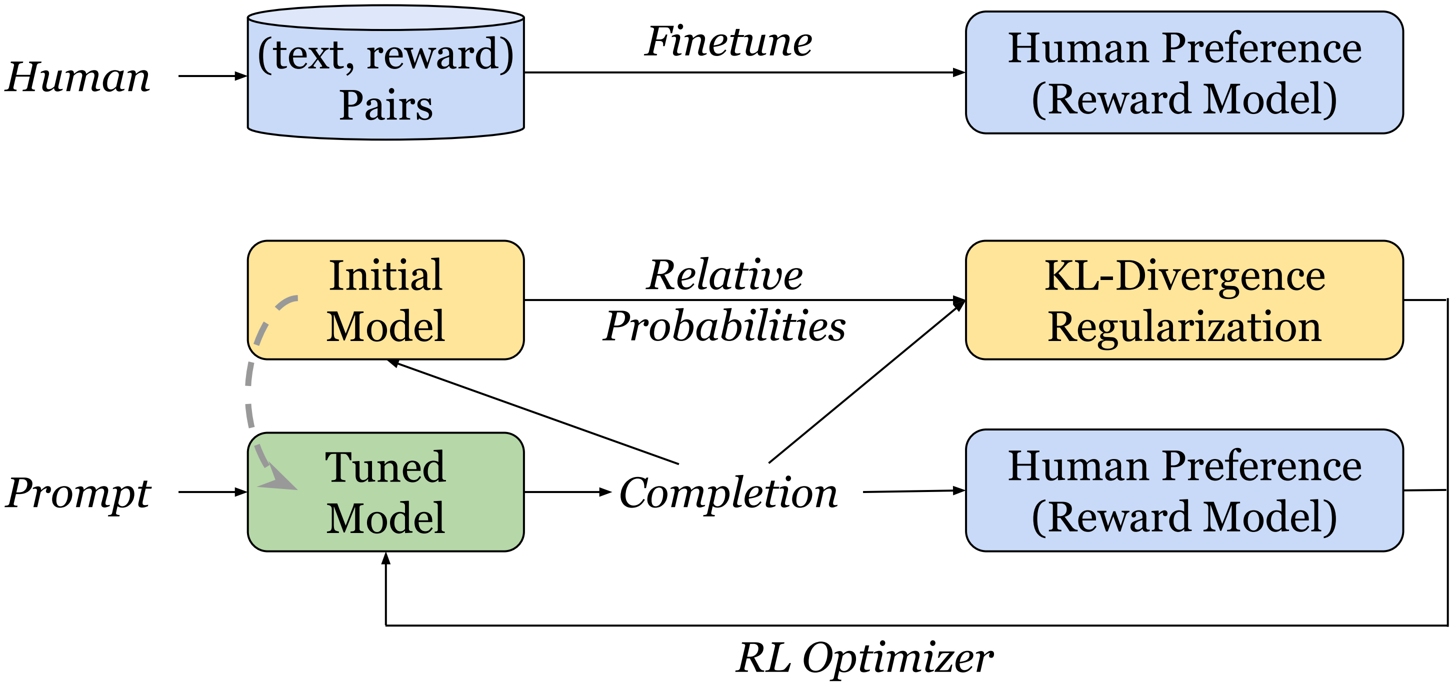
Corrective feedback in neural language models by trial and demo (TnD).

- Modeling corrective feedback in computational models with TnD.
  - The student model engages in *production-based learning*: to produce an initial utterance, followed by the teacher model generating its version of the text as a demonstration.
  - For the student model to recognize the teacher's response as preferable and to facilitate learning, these language outputs are evaluated by a reward function.



Method

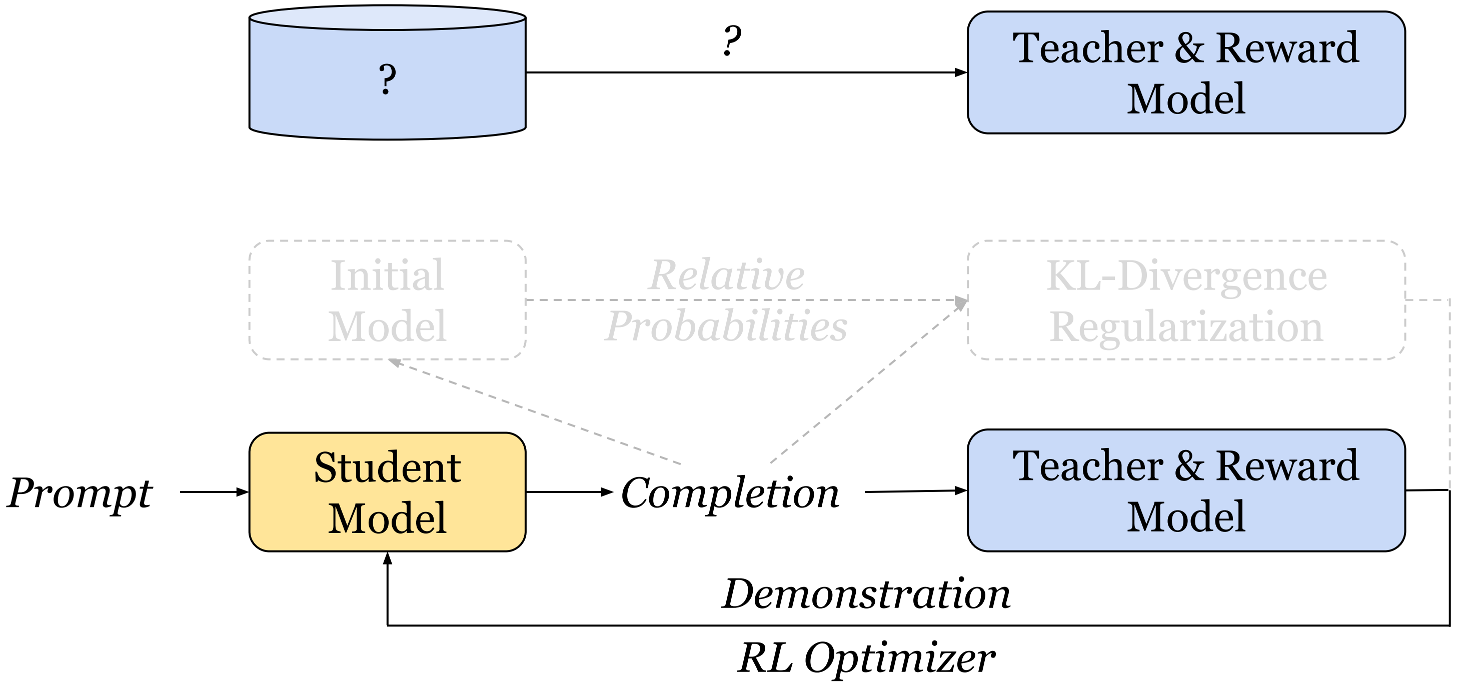
Reinforcement learning from human feedback (RLHF) in machine learning [16].



[16] Ouyang, Long, et al. "Training language models to follow instructions with human feedback." Advances in Neural Information Processing Systems 35 (2022): 27730-27744.

Method

Reinforcement learning for interactive language acquisition?



Method

Defining a reward in interactive language acquisition setting.

- RLHF algorithms enable neural language models to learn from extrinsic reward signals.
- But the interactive language acquisition scenario is different from this:
  - RLHF assumes a well-developed neural language model, and attempts to align it with human values and preferences.
  - RLHF requires an fixed external reward model developed from human scored text.
- The idea: The model's "age" can be a natural reward:
  - If a model @ 1,000 step produces a sentence that usually appear around 10,000 steps

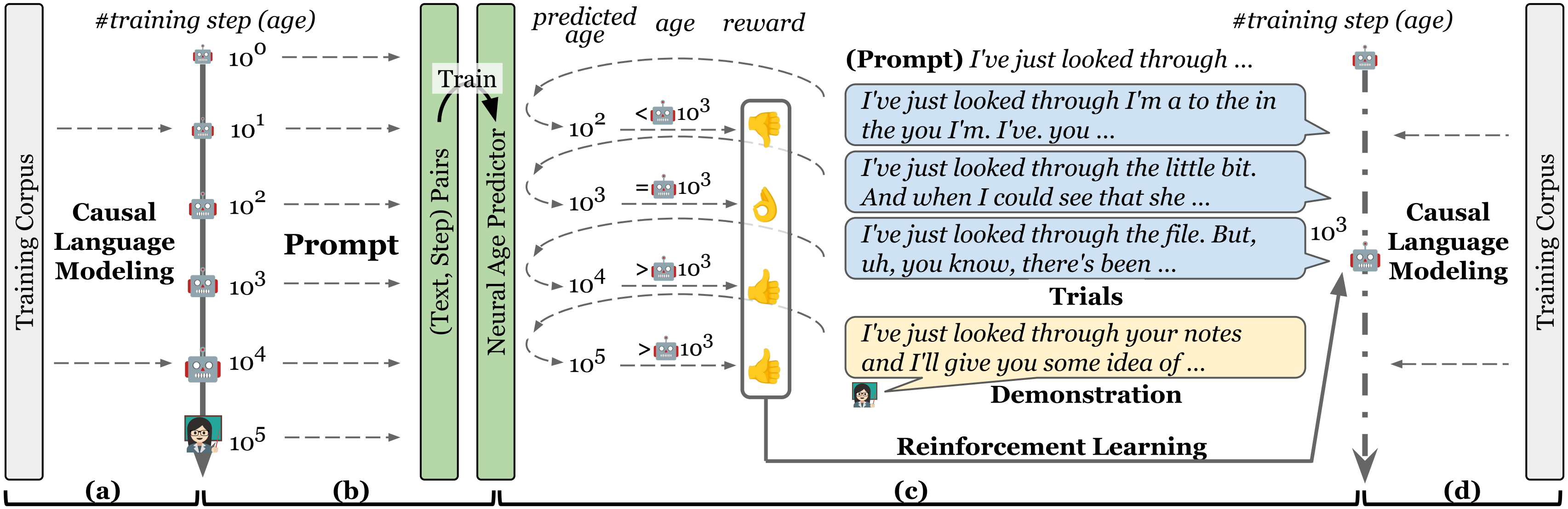
    -> **positive feedback;**
  - If a model @ 1,000 step produces a sentence that usually appear around 100 steps

    -> **negative feedback.**

Method

Babysitting a language model with student trials and teacher demonstrations.

- Trial-and-Demonstration (TnD) learning framework.



(a) Stage 1: Training a typical language model by causal language modeling.

(b) Stage 2: Training a neural age predictor from the trajectory of a typical language model.

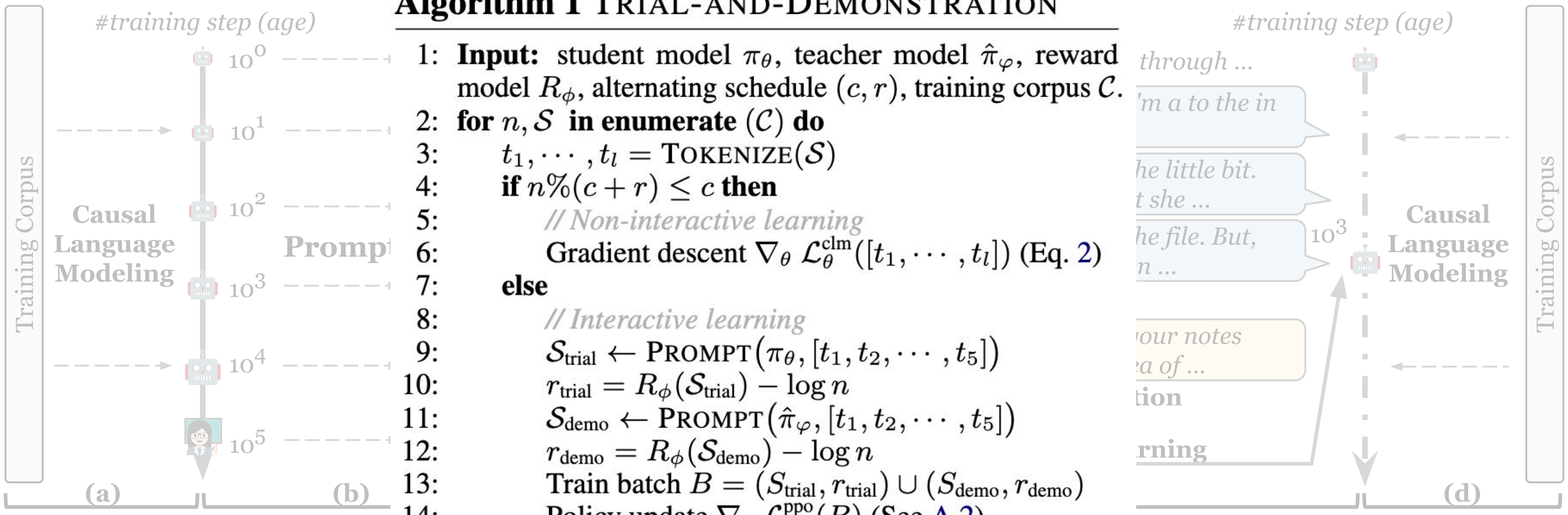
(c) Stage 3: The student interactively learn from trials and demonstrations by a pre-trained teacher, score by an age-conditioned reward.

(d) Alternating between interactive learning and non-interactive learning.

Method

Babysitting a language model with student trials and teacher demonstrations.

- Trial-and-Demonstration (TnD) learning framework.



Algorithm 1 TRIAL-AND-DEMONSTRATION

Input: student model π_θ, teacher model $\hat{\pi}_\varphi$, reward model R_ϕ, alternating schedule (c, r), training corpus $\mathcal{C}$.
for $n, \mathcal{S}$ **in enumerate** $(\mathcal{C})$ **do**
  $t_1, \cdots, t_l = \textsc{Tokenize}(\mathcal{S})$
  if $n\%(c+r) \le c$ **then**
    // Non-interactive learning
    Gradient descent $\nabla_\theta \mathcal{L}_\theta^{\text{clm}}([t_1, \cdots, t_l])$ (Eq. 2)
  else
    // Interactive learning
    $\mathcal{S}_{\text{trial}} \leftarrow \textsc{Prompt}(\pi_\theta, [t_1, t_2, \cdots, t_5])$
    $r_{\text{trial}} = R_\phi(\mathcal{S}_{\text{trial}}) - \log n$
    $\mathcal{S}_{\text{demo}} \leftarrow \textsc{Prompt}(\hat{\pi}_\varphi, [t_1, t_2, \cdots, t_5])$
    $r_{\text{demo}} = R_\phi(\mathcal{S}_{\text{demo}}) - \log n$
    Train batch $B = (\mathcal{S}_{\text{trial}}, r_{\text{trial}}) \cup (\mathcal{S}_{\text{demo}}, r_{\text{demo}})$
    Policy update $\nabla_\theta \mathcal{L}_\theta^{\text{ppo}}(B)$ (See A.2)
  end if
end for
Output: π_θ

(a) Stage 1: Training a typical language model by causal language modeling.

(d) Alternating between interactive learning and non-interactive learning.

Method

Babysitting a language model with student trials and teacher demonstrations.

- Trial-and-Demonstration (TnD) learning framework.

Student Model

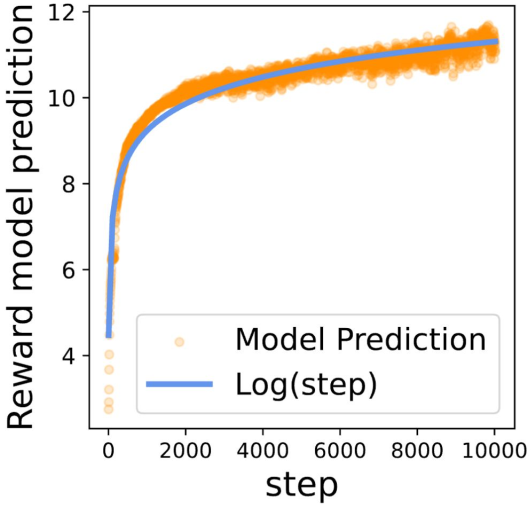
GPT-2
d=768

Teacher Model

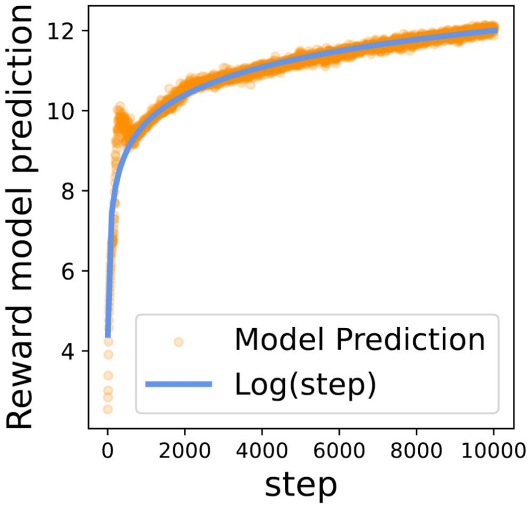
GPT-2
d=768

Reward Model

LLaMA-2-7B



(a) BookCorpus dataset.

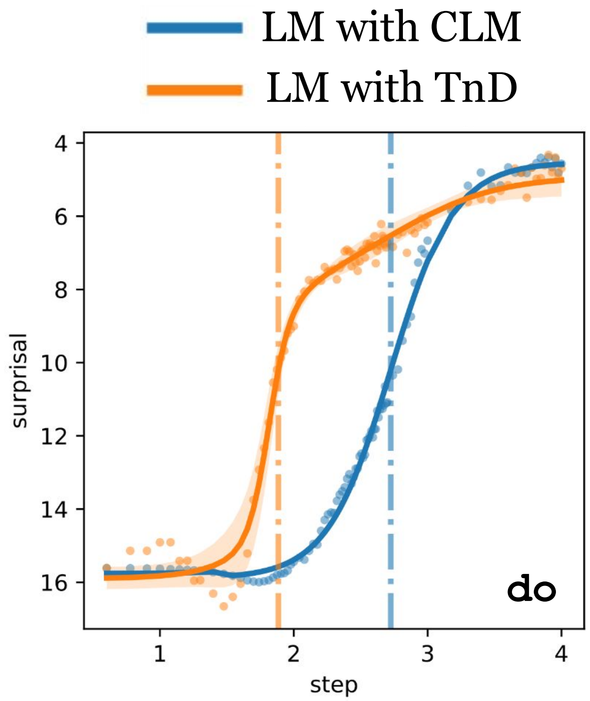


(b) BabyLM dataset.

Evaluation

Evaluating how well and how fast a neural language model acquires a word.

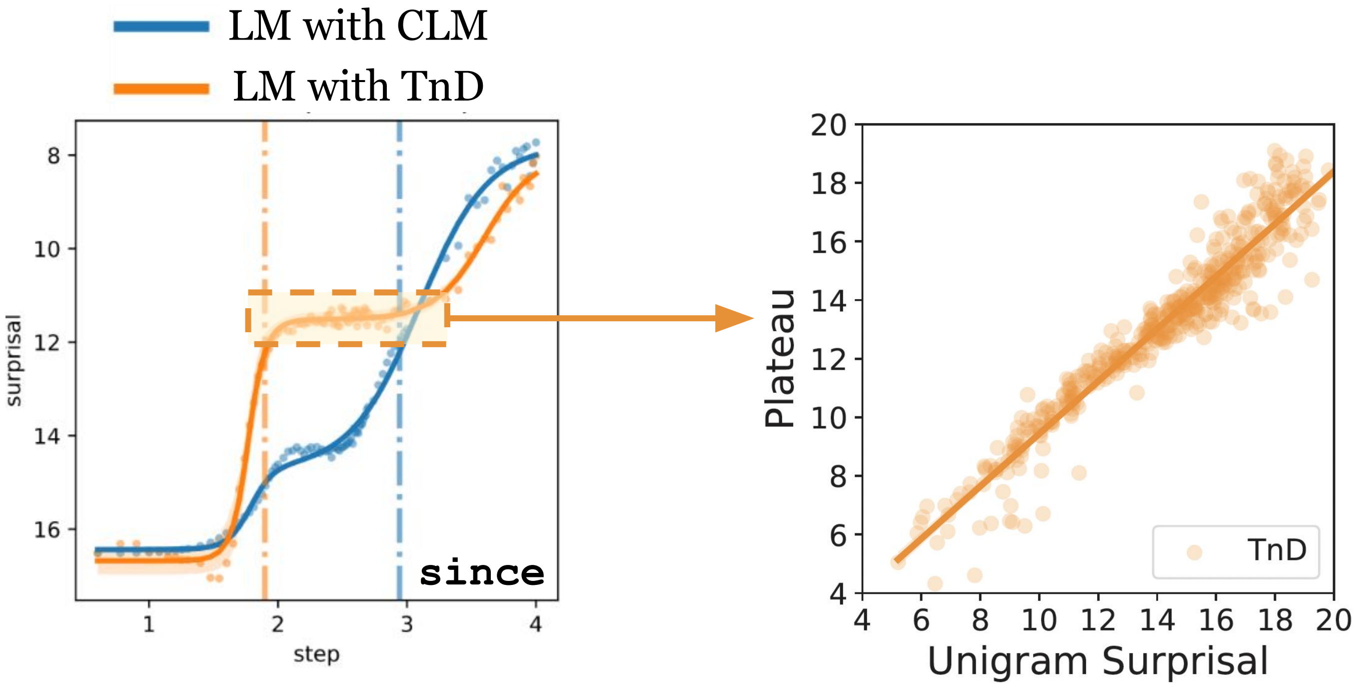
- Surprisal / Perplexity:
  - Measure of how unexpected that particular word is in a given context.
  - For a word *w* with probability *P(w)* given a certain context, its surprisal is calculated as *-log(P(w))*.
  - Averaged ≥100 context where the word is used.



Evaluation

Evaluating how well and how fast a neural language model acquires a word.

- Learning curve fitting [9].

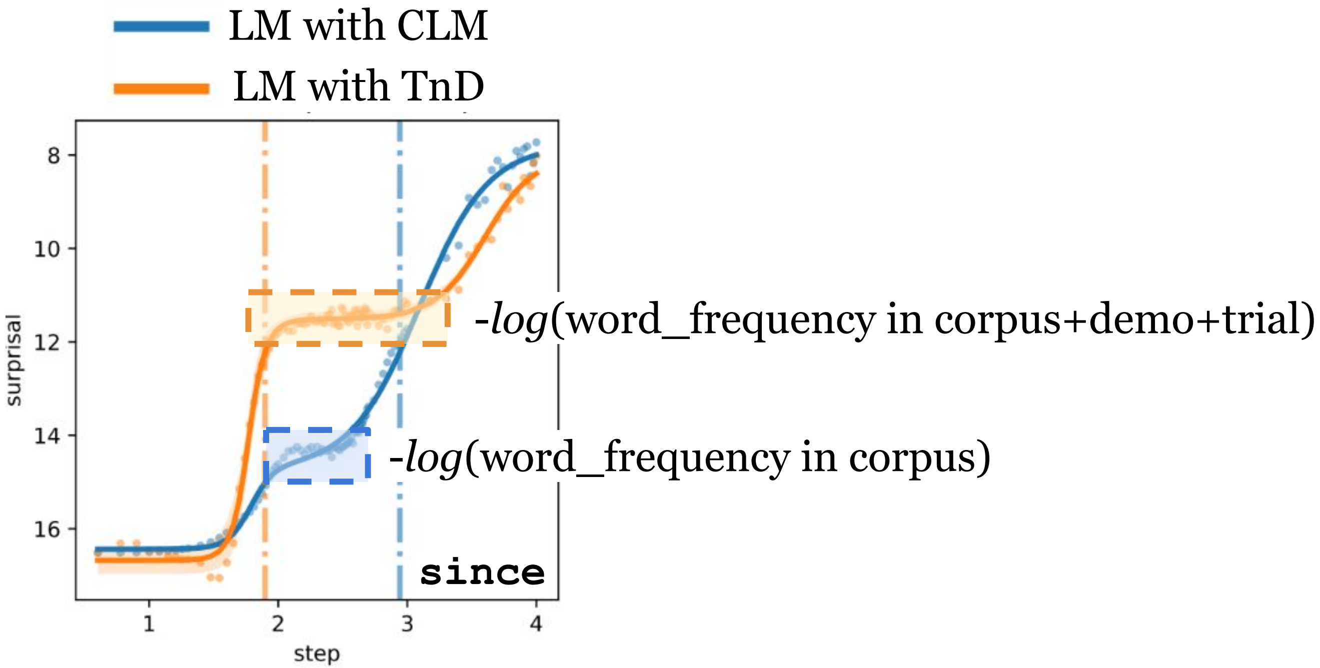


[9] Tyler A. Chang and Benjamin K. Bergen. Word acquisition in neural language models. Transactions of the Association for Computational Linguistics 10 (2022): 1-16.

Evaluation

Evaluating how well and how fast a neural language model acquires a word.

- Learning curve fitting [9].

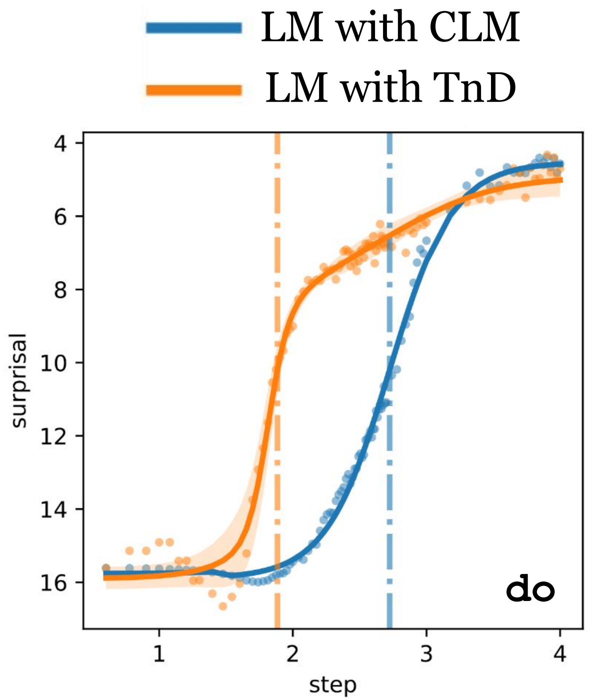


[9] Tyler A. Chang and Benjamin K. Bergen. Word acquisition in neural language models. Transactions of the Association for Computational Linguistics 10 (2022): 1-16.

Evaluation

Evaluating how well and how fast a neural language model acquires a word.

- Age of Acquisition (AoA):
  - Measure of how *fast* a model learns a word.
  - We established a cutoff surprisal where we considered a given word "learned."
  - In child word acquisition studies, an analogous cutoff is established when 50% of children produce a word [17].
  - [9] determined our cutoff to be 50% between a baseline surprisal (predicting words based on random chance) and the minimum surprisal.
  - We further evaluate on cutoffs in 50%-95% with step size 5%.



[9] Tyler A. Chang and Benjamin K. Bergen. Word acquisition in neural language models. Transactions of the Association for Computational Linguistics 10 (2022): 1-16.
[17] Mika Braginsky, Daniel Yurovsky, Virginia Marchman, and Michael Frank. 2016. From uh-oh to tomorrow: Predicting age of acquisition for early words across languages. In Proceedings of the Annual Meeting of the Cognitive Science Society.

Experiments

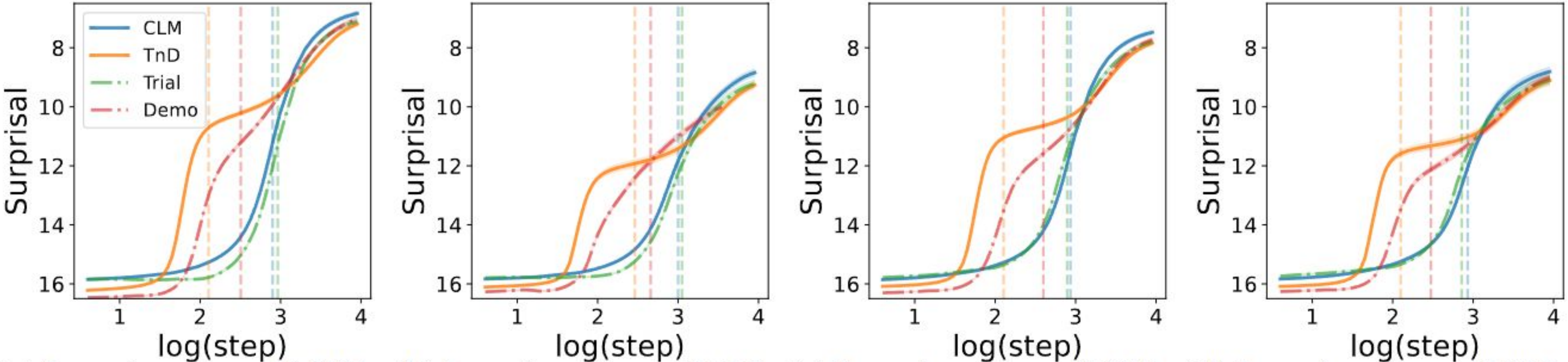
Experiment setups.

- Training Corpus:
  - Corpus 1: BookCorpus (Commonly used in AI/NLP);
  - Corpus 2: BabyLM Challenge (Cognitively plausible, CHILDES, Subtitles, BNC, TED talks, children's books);
- Test vocabulary
  - Set 1: Common Words
  - Set 2: Communicative Development Inventories (CDIs) Words
- Controlled studies
  - 5 random seeds for each model, systematic search of learning rates and alt. frequency.

Results and Discussions

Main Findings.

- TnD leads to faster acquisition, but eventually the final performance converges;
- Both trial and demo are important!

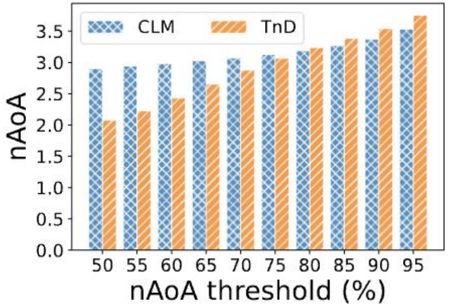


(a) Learning curve of CMN words on BookCorpus. (b) Learning curve of CDI words on BookCorpus. (c) Learning curve of CMN words on BabyLM. (d) Learning curve of CDI words on BabyLM.

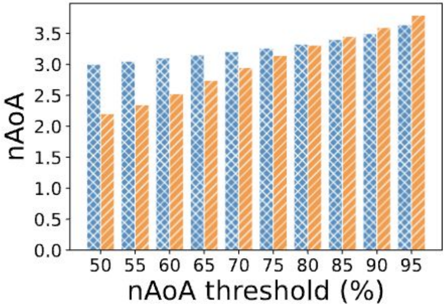
Results and Discussions

Main Findings.

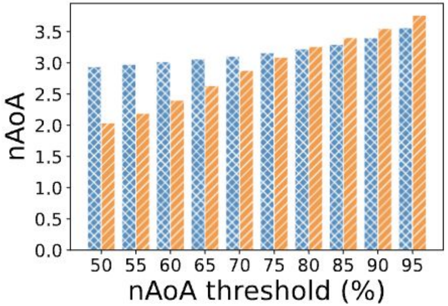
- TnD leads to earlier neural age of acquisition (nAoA) when cutoff < 80%;
- The rapid acquisition stage diminished in the later training stage.



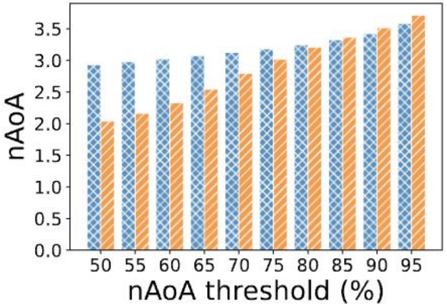
(a) nAoA at each cutoff of CMN words on BookCorpus.



(b) nAoA at each cutoff of CDI words on BookCorpus.



(c) nAoA at each cutoff of CMN words on BabyLM.

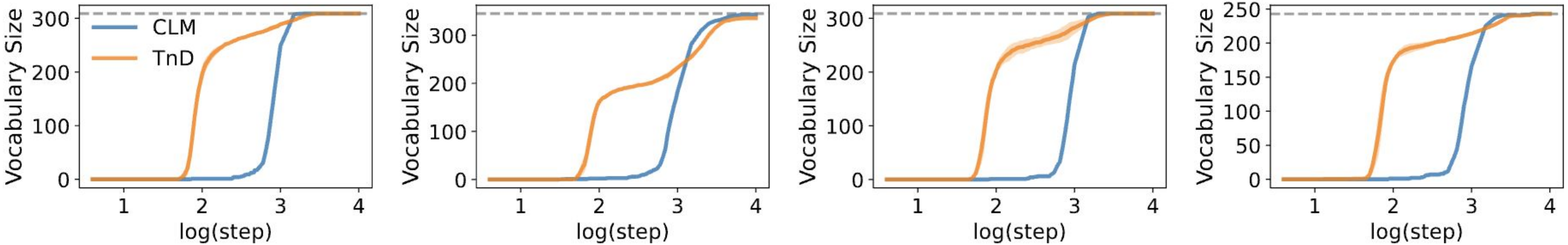


(d) nAoA at each cutoff of CDI words on BabyLM.

Results and Discussions

Main Findings.

- Students under the TnD framework quickly picked up a large volume of effective vocabulary, but eventually their vocabulary capacities have converged to the CLM baseline as expected.



(a) Effective vocabulary size in CMN words on BookCorpus. (b) Effective vocabulary size in CDI words on BookCorpus. (c) Effective vocabulary size in CMN words on BabyLM. (d) Effective vocabulary size in CDI words on BabyLM.

Results and Discussions

Further Finding 1:

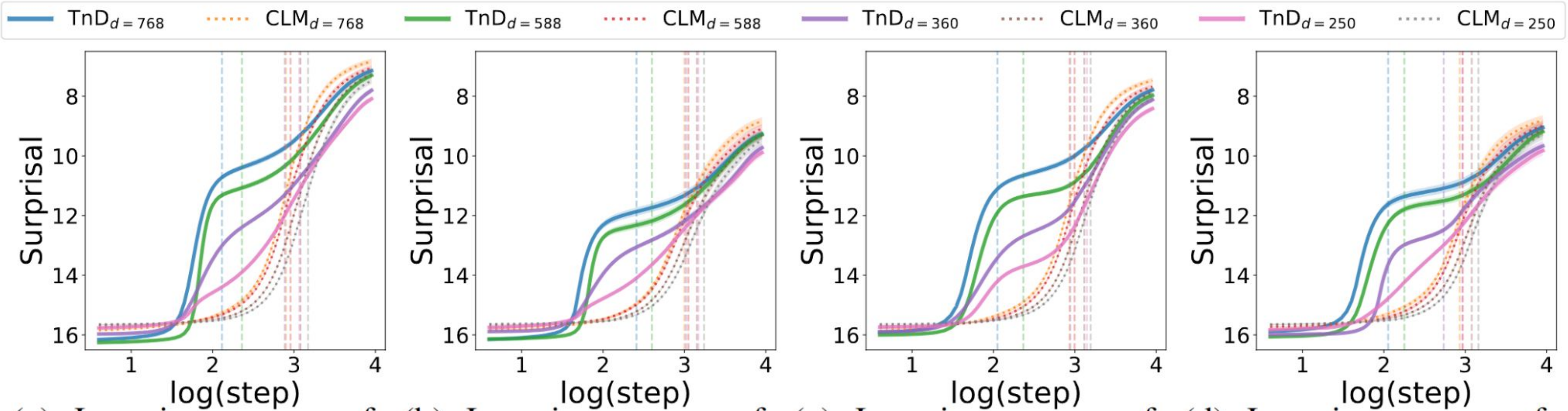
- The original student GPT-2 model has a dimension of d=768 (12 attention heads each with a dimension of 64).
- We now keep all experimental setups untouched but smaller student models with dimensions of 588 (12 x 49), 360 (10 x 36), and 250 (10 x 25) respectively.

Teacher Model	Student Model
GPT-2 **d=768**	**GPT-2** **d=768/588/360/250**

Results and Discussions

Further Finding 1:

- Each TnD model outperforms the CLM baseline of the same size, and even surpasses CLM baselines of large capacity in early steps.



(a) Learning curve sof smaller models of CMN words on BookCorpus.

(b) Learning curve sof smaller models of CDI words on BookCorpus.

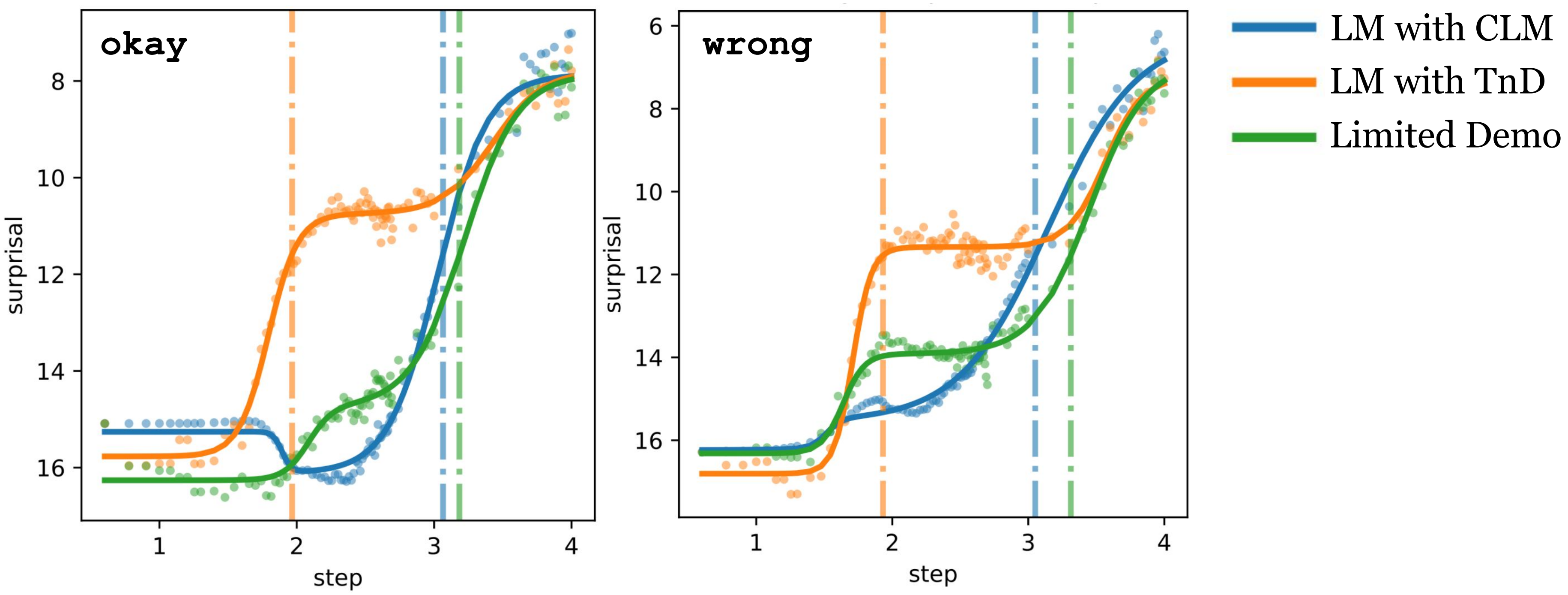
(c) Learning curve sof smaller models of CMN words on BabyLM.

(d) Learning curve sof smaller models of CDI words on BabyLM.

Results and Discussions

Further Finding 2:

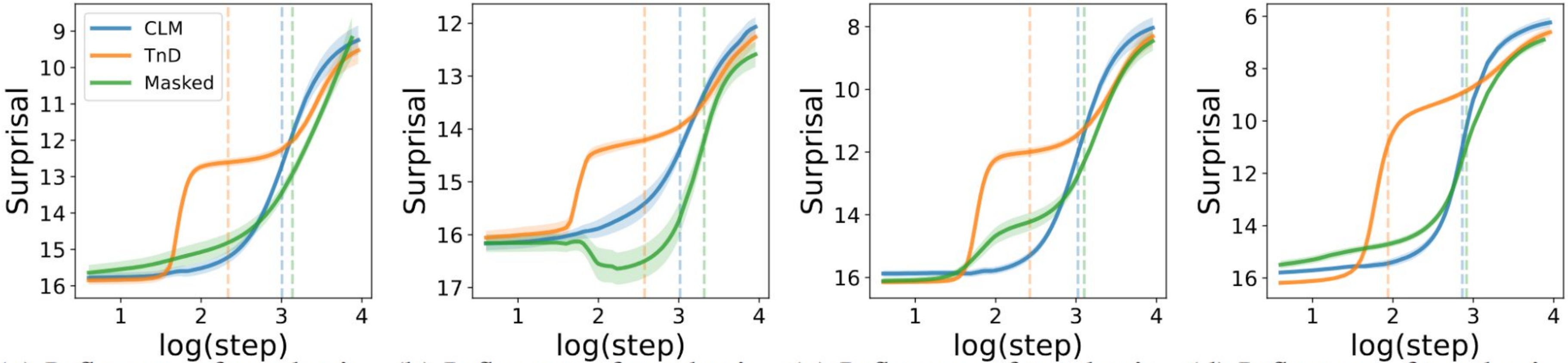
- Teacher's word preferences affect student's word-specific acquisition.



Results and Discussions

Further Finding 2:

- Teacher's word preferences affect student's word-specific acquisition.



(a) Influence of teacher's word preferences in CMN words on BabyLM.

(b) Influence of teacher's word preferences in CDI words on BabyLM.

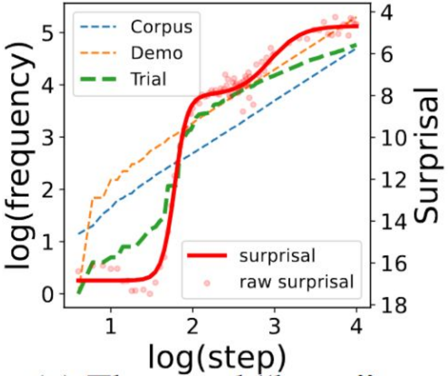
(c) Influence of teacher's word preferences in CMN words on BookCorpus.

(d) Influence of teacher's word preferences in CDI words on BookCorpus.

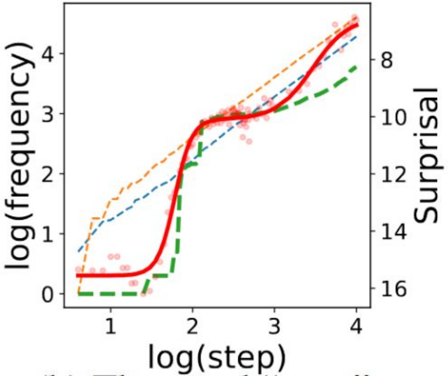
Results and Discussions

Further Finding 3:

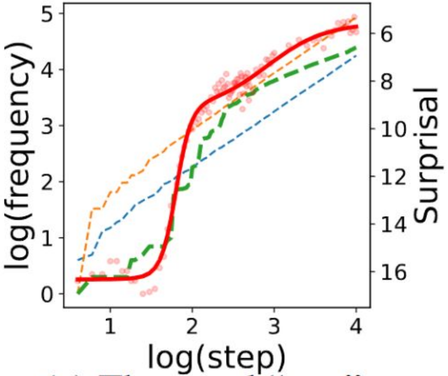
- Attempts of trials are highly correlated to the acquisition of functional words and predicates, whose semantics depend on other words …



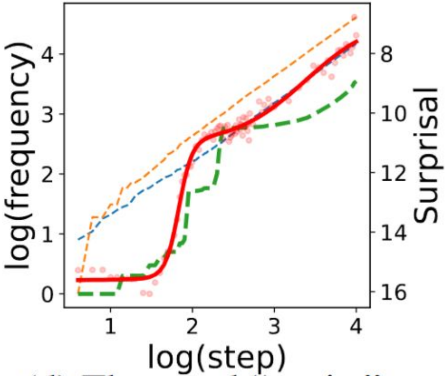
(a) The word "have."



(b) The word "now."



(c) The word "see."

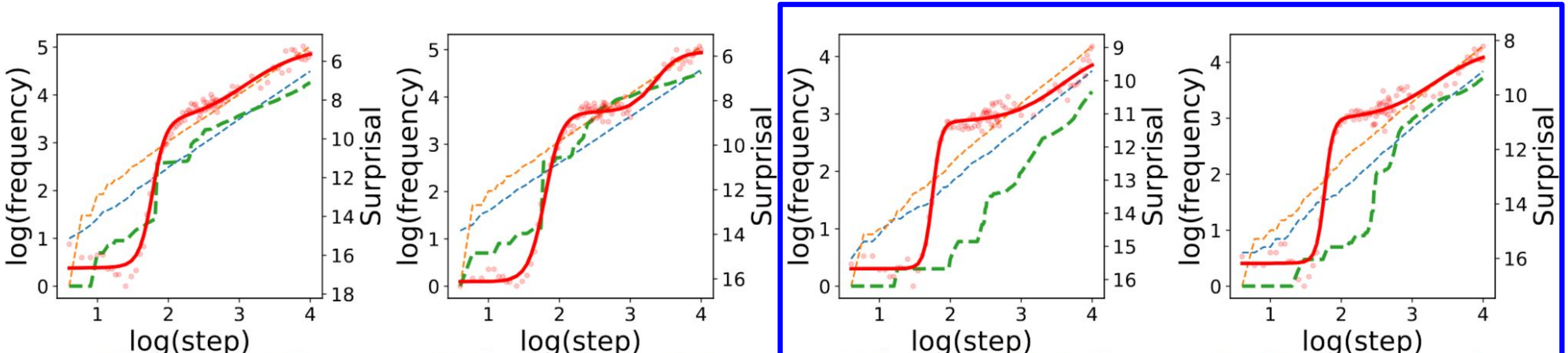


(d) The word "again."

Results and Discussions

Further Finding 3:

- Attempts of trials are highly correlated to the acquisition of functional words and predicates, whose semantics depend on other words, **but not groundable concrete nouns, whose semantics depend on the grounded objects**.



(e) The word "if." (f) The word "back." (g) The word "light." (h) The word "car."

Results and Discussions

Further Finding 3:

- We present the β weights and Pearson correlation r between their mean surprisal and cumulative word frequency over the course of training from predictor analysis.

POS	Freq.	BabyLM Corpus				BookCorpus			
		CMN		CDI		CMN		CDI	
		β	r	β	r	β	r	β	r
noun	trial	-0.36	-0.90	-0.25	-0.85	-0.38	-0.92	-0.31	-0.85
	demo	**-0.67**	-0.93	**-0.73**	-0.89	**-0.56**	-0.93	**-0.67**	-0.87
	corpus	-0.51	-0.93	-0.53	-0.88	**-0.56**	-0.93	-0.60	-0.87
pred	trial	**-0.70**	-0.90	**-0.72**	-0.86	**-0.49**	-0.92	**-0.54**	-0.88
	demo	-0.33	-0.93	-0.30	-0.90	**-0.49**	-0.92	-0.45	-0.90
	corpus	-0.22	-0.93	-0.19	-0.90	-0.45	-0.93	-0.40	-0.87
func	trial	**-0.67**	-0.93	**-0.72**	-0.92	**-0.67**	-0.94	**-0.59**	-0.93
	demo	-0.39	-0.92	-0.21	-0.90	-0.22	-0.91	-0.37	-0.87
	corpus	-0.17	-0.92	-0.25	-0.91	-0.35	-0.92	-0.35	-0.90

Limitations and Future Work

Limitations

- **Iterative setting:** This experiment can be conducted iteratively by replacing the teacher model with the student model from previous iterations.
- **Reward model**: We employ a robust language model (LLaMA-2-7B) as a reward model to concentrate on the roles of trials and demonstrations without concerns about reward quality. Future research should explore the impact of using less accurate reward models.
- **Intrinsic Reward:** child might instinctively feel satisfaction from producing a sound that echoes a meaningful memory [18].
- **Tokenizer:** the reliance on the BPE tokenizer shared some linguistic priors and fail to capture the early language elements in child such as sounds effects and animal sounds.
- **Language:** expend on other languages.

[18] Thorndike, Edward. "Animal intelligence; experimental studies." Animal behavior series (1911).

Thanks!