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

Ziqiao Ma*, Zekun Wang*, Joyce Chai
Computer Science and Engineering, University of Michigan

Motivations

- Humans are social beings, and we learn language from interactions;
- A critical component of social interactions that language grounds to is the feedback provided by the caregivers;
- We study the role of corrective feedback in neural language learning through controlled computational experiments;
- Through ablation studies, our models can serve as proof of concept to verify mechanisms that are practically effective for machines, and generate hypotheses that are possible for cognitive learners.

Key Findings (TL;DR)

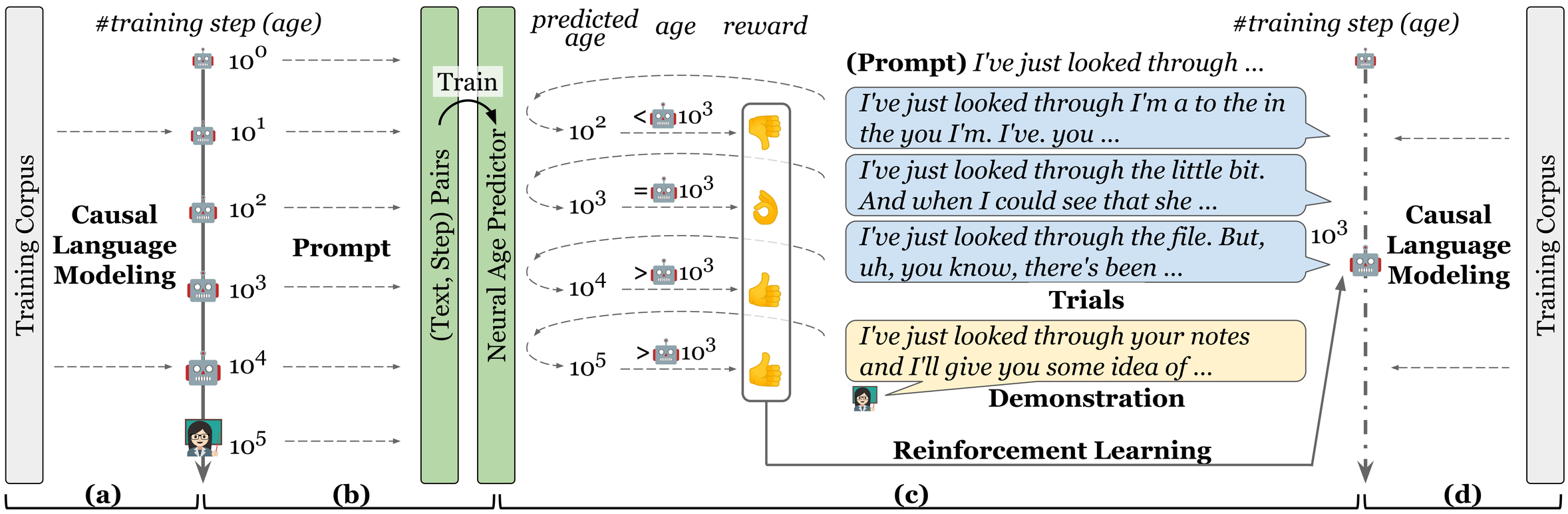
- Training language models from scratch with interactive language modeling by trials and demonstrations (TnD);
- TnD accelerates word acquisition for student models of equal and smaller numbers of parameters. Both trials and demonstrations matter;
- Teacher's choices of words influence students' word-specific learning efficiency;
- Students demonstrate a practice-makes-perfect effect, evident by a strong correlation between the word frequency trials and their respective learning curves.

Method

The trial and demonstration (TnD) learning framework with 3 components: student trials, teacher demonstrations, and an "age-conditioned" reward conditioned on language competence over time.

- Students do production-based learning: to produce an initial utterance, followed by the teacher model generating its version of the text as a demonstration.
- We use a neural age predictor to estimate the expected training step $\hat{n}$ when an utterance typically emerges, then normalized by the actual training step n.

Method Illustrated: Interactive Language Learning by Trials and Demos



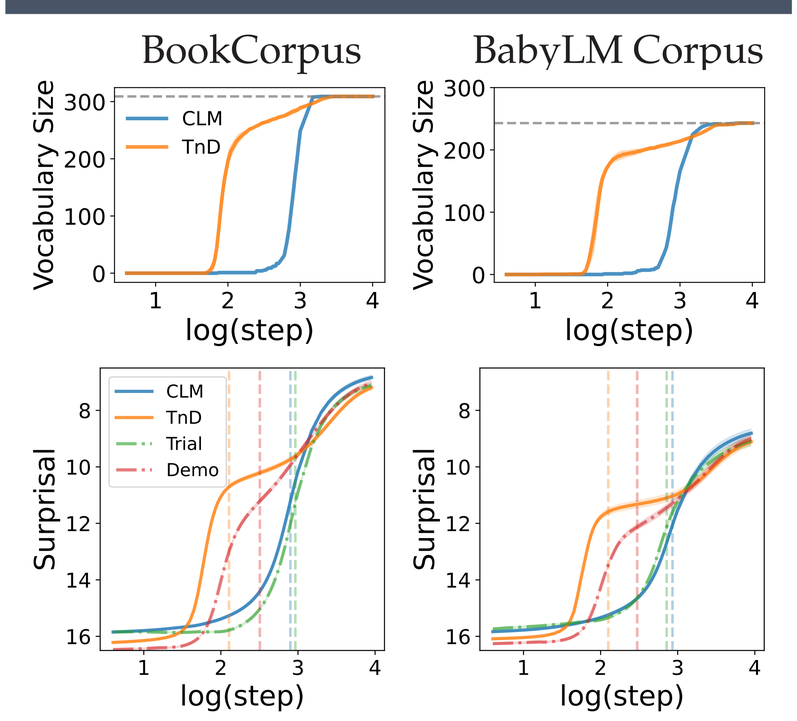
(a) Stage 1: Training a typical language model by causal language modeling.

(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.

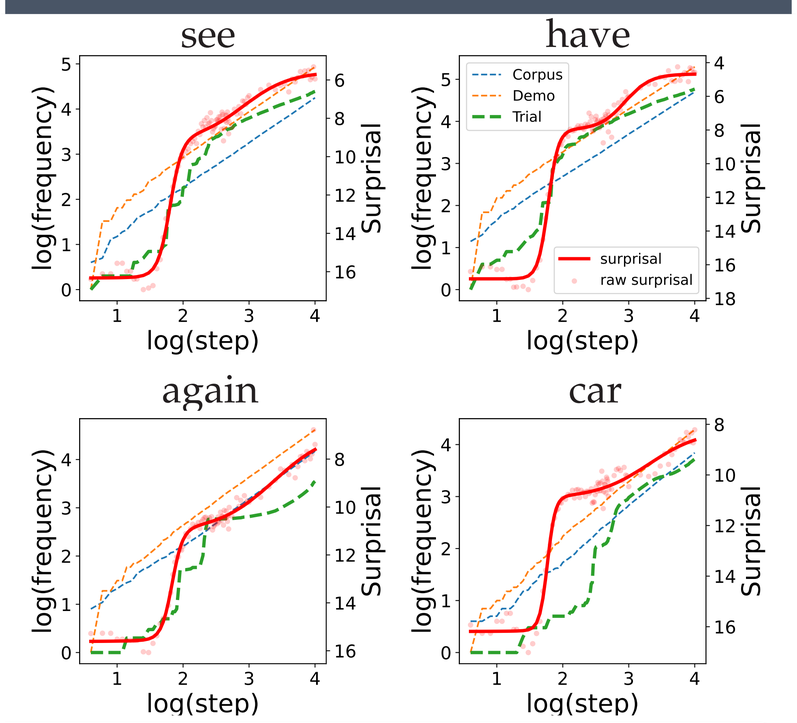
Main Results



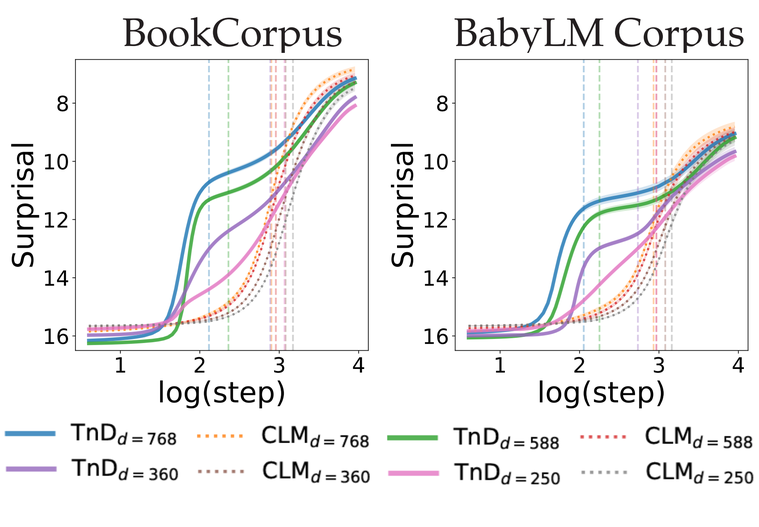
Metrics

- Surprisal and learning curves. For each occurrence of a word w, the surprisal is $-\log_2 P(w)$, then averaged over all occurrences. We adopt a double-sigmoid function to fit the learning curve.
- Neural age of acquisition (nAoA@$n\%$). The training step at which a surprisal cutoff of $n\%$ between the minimum and maximum surprisal levels is achieved.
- Vocabulary size. The effective vocabulary size relative to a test set of vocabulary. A word is deemed acquired at step n if $\texttt{nAoA@0.50} \leq n$.

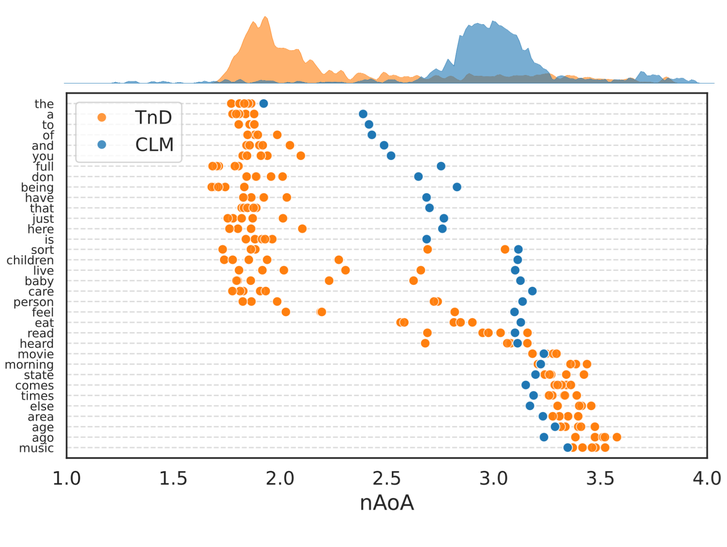
Student Effect



Distillation



Trajectory



Teacher Effect