Cambridge ALTA
Institute for Automated Language Teaching and Assessment

UNIVERSITY OF CAMBRIDGE

Bilingual Small Language Models as Cognitive Proxies for LLM Interaction and Calibration

Suchir Salhan
Department of Computer Science & Technology, University of Cambridge, U.K. `sas245@cam.ac.uk`

Small Language Models (SLMs): Goals and Rationale

Small Language Models (SLMs), typically under 1B parameters, provide valuable, interpretable, and efficient alternatives to large models. They are especially suitable for proprietary, task-specialised applications such as query routing in chatbot systems or edge/on-device learning.

My PhD Work: SLMs as *Learner Models*

Problem: High-precision learner representations are essential for personalised and adaptive learning and assessment.
Solution: I propose bilingual SLMs for second language adaptation — or **L2LMs** — which simulate the developmental trajectories of second-language learners with a typologically-diverse L1s. Scales and extends work in the BabyLM Shared Task which train SLMs on around 100M tokens of Child-Directed Speech and Simplified Texts to simulate the *volume* and *nature* of linguistic input provided to humans in first language acquisition [1,2,5].

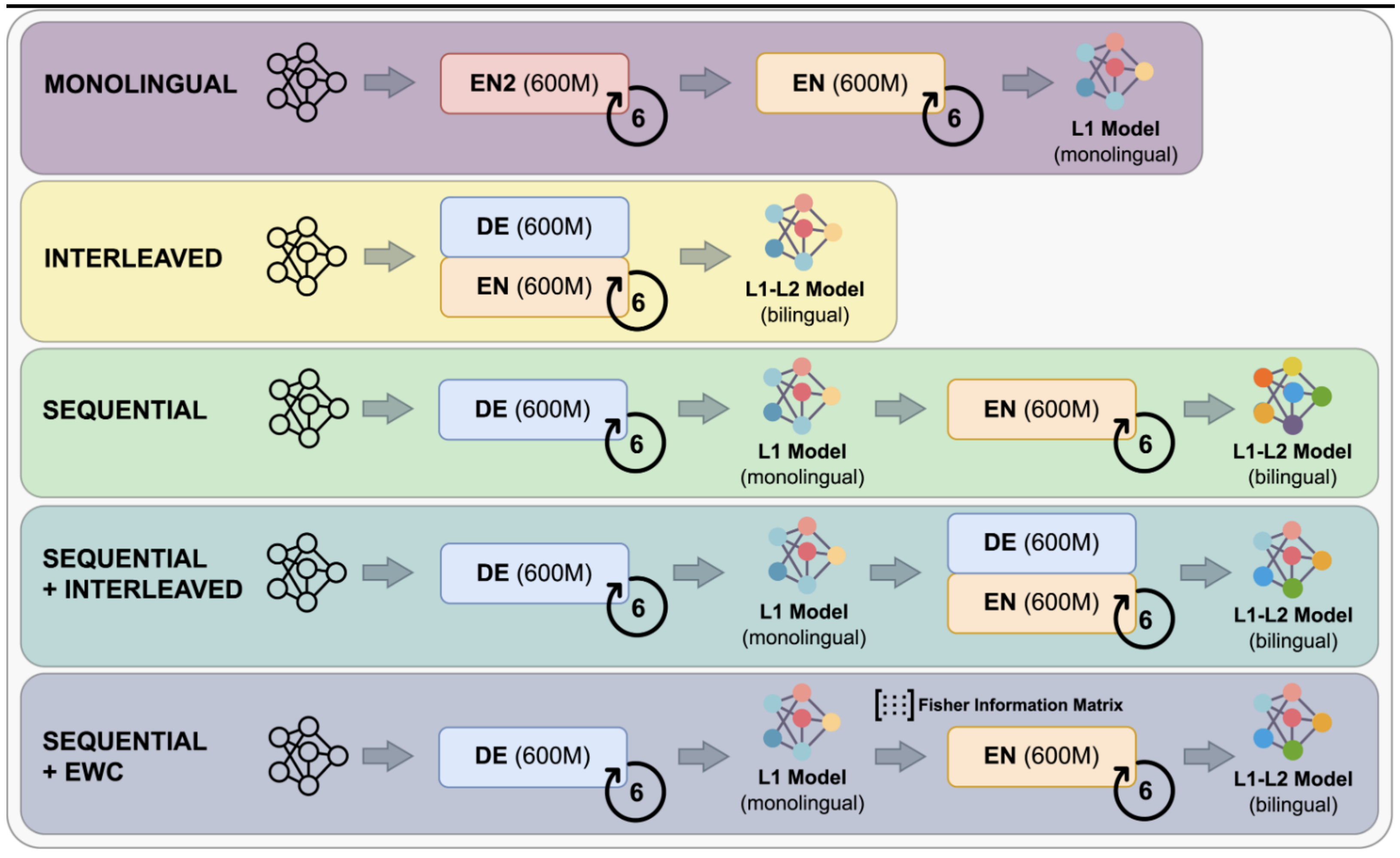
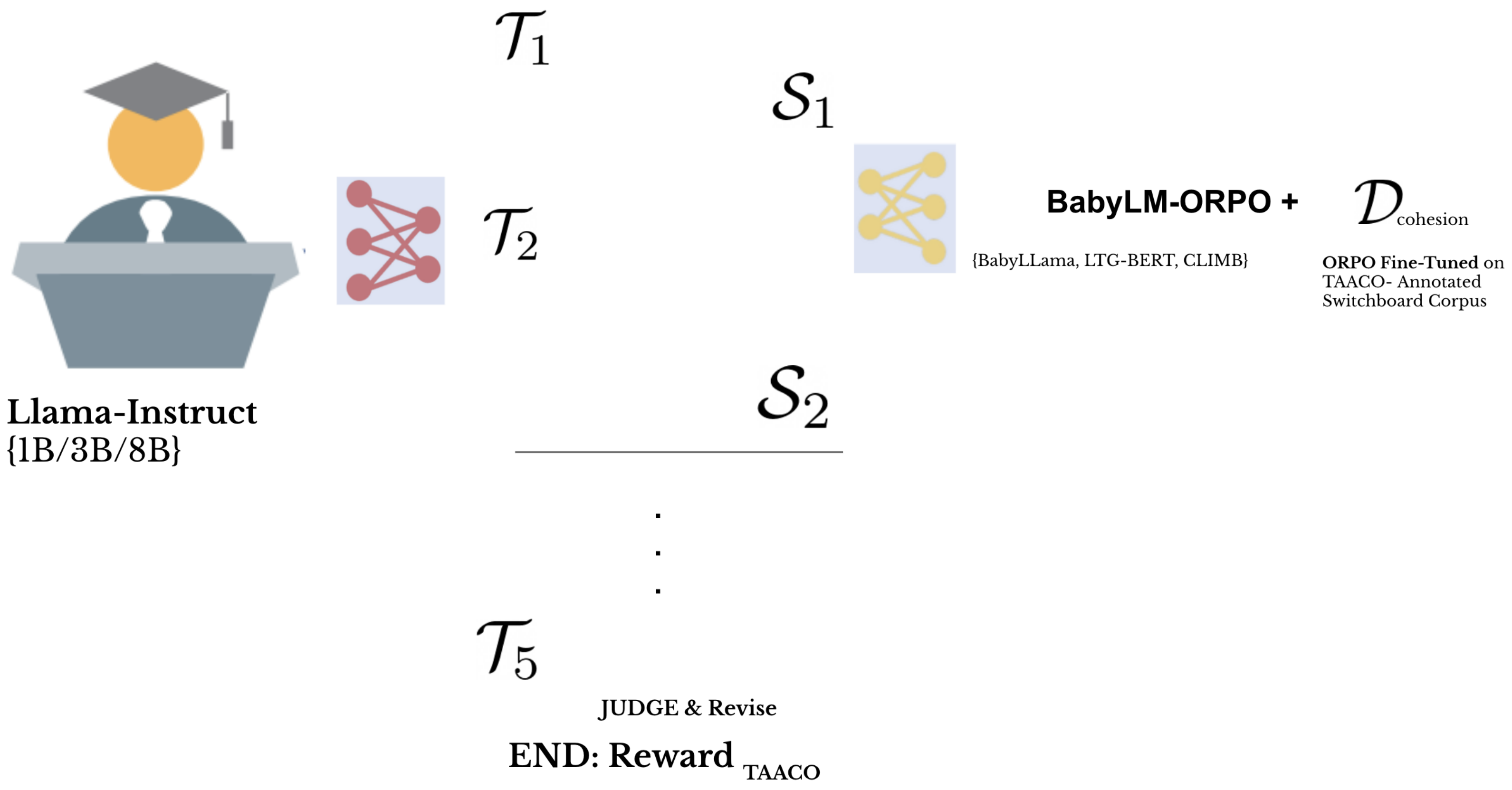


Figure 1:Conceptual Architecture of L2LMs and Sequential/Simultaneous Bilingual SLMs from Constantinescu et al (2025)

Aims: Rich Learning Dynamics and Checkpointing of L2LMs for enhanced interpretability, building on the Pico Framework [3]. Prioritise **developmental interpretability** to see exactly where models begin to represent certain linguistic features, how representations stabilise (or oscillate) over training, and which parts of the network converge faster than others to develop more cognitively-aligned pretraining strategies. Model design to reflect the *diversity of L1 backgrounds* (over 55 distinct L1 language families in the Cambridge Learner Corpus)

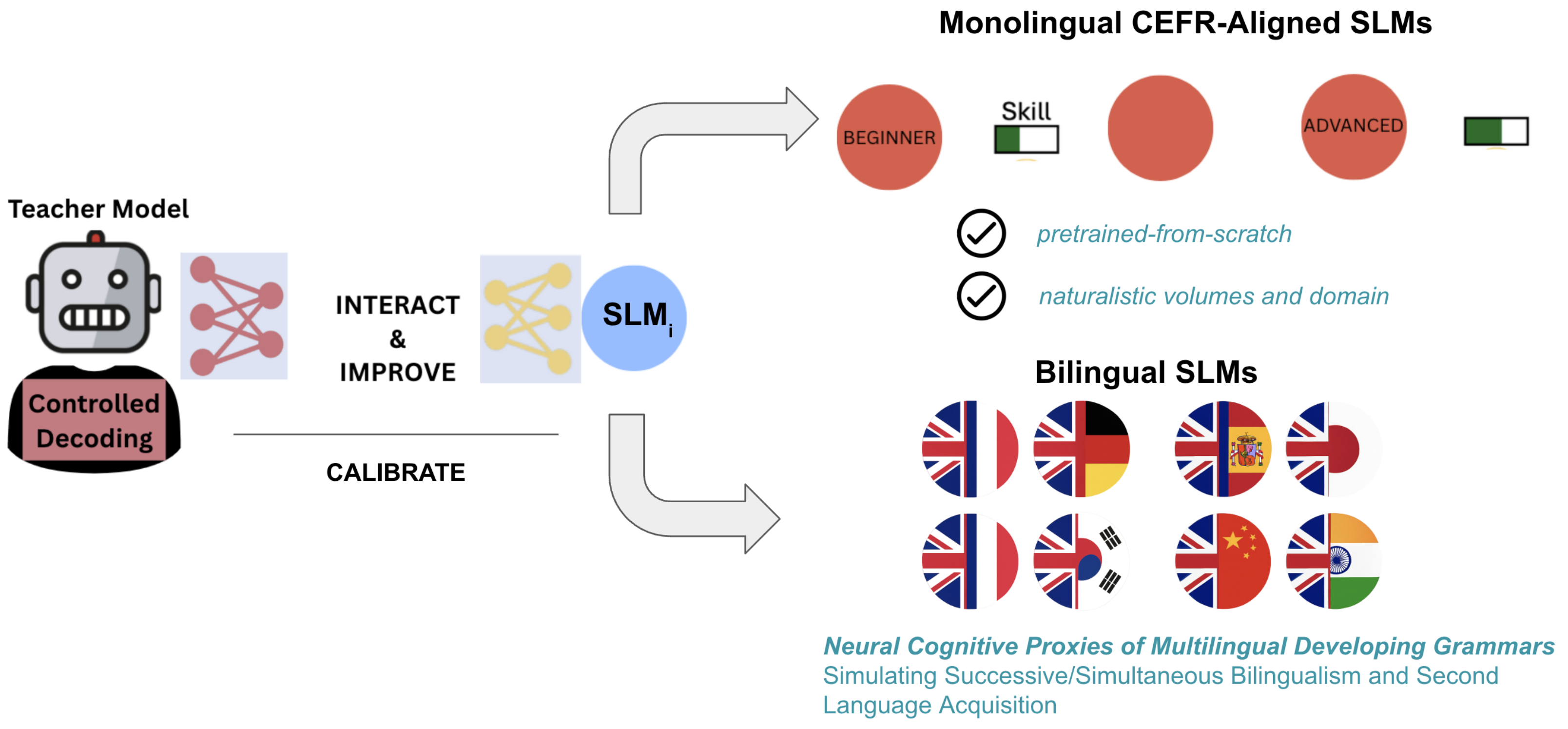
SLMs and LLM Interaction

Preliminary Work for BabyLM Interaction Track [6]: Provide rewards from a Teacher LLM to reward BabyLMs to produce *more cohesive long-form generations*. BabyLMs trained on 100M (the Strict Track) are our Student Model that interact with a Llama-Instruct 1B/3B/8B parameter model (the Teacher LLM) in a multi-turn dialogue. Dialogue is initiated by the Teacher in $\mathcal{T}_1$, prompting the student BabyLM for a continuation of a prompt in utterance $\mathcal{S}_1$.


$\mathcal{T}_1$
$\mathcal{S}_1$
$\mathcal{T}_2$
BabyLM-ORPO + $\mathcal{D}_{\text{cohesion}}$
{BabyLLama, LTG-BERT, CLIMB}
ORPO Fine-Tuned on TAACO- Annotated Switchboard Corpus
Llama-Instruct {1B/3B/8B}
$\mathcal{S}_2$
$\mathcal{T}_5$
JUDGE & Revise
END: Reward $_{\text{TAACO}}$

Next Steps:

Figure 2:Proposed Framework for **LLM-L2LM Interaction** and **LLM Interaction with CEFR-Aligned LLMs**


Monolingual CEFR-Aligned SLMs
BEGINNER
Skill
ADVANCED
Teacher Model
Controlled Decoding
INTERACT & IMPROVE
SLM$_i$
CALIBRATE
✓ pretrained-from-scratch
✓ naturalistic volumes and domain
Bilingual SLMs
Neural Cognitive Proxies of Multilingual Developing Grammars
Simulating Successive/Simultaneous Bilingualism and Second Language Acquisition

Selected References

[1] **Salhan, S.A** (2025) **Linguistics in the Age of Language Models: What can Cognitively-Inspired Language Models offer to Linguistic Theory?** (Position Paper in *Cambridge Occasional Papers in Linguistics (CoPiL)*, Accepted, Volume 17) [2] Arnett, C. On the Acquisition of Shared Grammatical Representations in Bilingual Language Models *in press* [3] Diehl Martinez, R., Demitri Africa, D., Weiss, Y., **Salhan, S.A.**, Daniels, R., & Buttery. P.J. (2025) **Pico: A Lightweight Framework for Studying Language Model Learning Dynamics** (*under review*) [4] Goriely, Z., **Salhan, S.A.**, Lesci, P.,Cheng. J., & Buttery. P.J. (2025) **ByteSpan: Information-Driven Subword Tokenisation** (Accepted ICML 2025 Tokenisation Workshop (TokShop, Non-Archival)) [5] **Salhan, S.A.**, Diehl Martinez, R., Goriely, Z., & Buttery. P.J. (2024) **Less is More: Pre-Training Cross-Lingual Small-Scale Language Models with Cognitively-Plausible Curriculum Learning Strategies** (Accepted Paper @ BabyLM Workshop in EMNLP 2024). [6] **Salhan, S.A.**, Galvan-Sosa, D., Rooein, D., Gao, H., Gaudeau, G., Yuan, Z., Caines, A.P. & Buttery, P.J (2025) *Talking BabyLMs* (submitted for BabyLM 2025, interaction Track)

This work is supported by Cambridge University Press & Assessment