Adaptive Language Learning Quiz for Newcomer 6th Graders

JORGE SANDOVAL, UC Berkeley, United States of America

This paper presents the development and preliminary evaluation of an adaptive language learning quiz tailored for newcomer 6th-grade English language learners. Grounded in Vygotsky's Zone of Proximal Development (ZPD) and the principles of scaffolding, this system dynamically adjusts task difficulty in response to learner performance. Utilizing the CommonLit Ease of Readability (CLEAR) Corpus as a data source, we employed a fine-tuned T5 model to generate a diverse array of tasks—fill-in-the-blank, synonym matching, and sentence rearrangement—aligned with 6th-grade proficiency levels. Baseline experiments with a BERT-based approach established initial benchmarks, while subsequent evaluations of the adaptive T5-based model demonstrated improved accuracy, BLEU, and ROUGE scores. Although the current prototype did not consistently produce "just right" challenges to maximize learning gains, it offers a promising starting point. The system's design and outcomes underscore the potential of adaptive technologies to address linguistic diversity in the classroom, as well as the need for iterative refinement. Future directions include integrating more nuanced adaptation strategies, enhancing task quality, and conducting longitudinal studies to measure long-term impacts on language acquisition and academic integration.

CCS Concepts: • **Applied computing → Education**.

Additional Key Words and Phrases: Adaptive Learning, Language Acquisition, Educational Technology, Scaffolding

ACM Reference Format:
Jorge Sandoval. 2024. Adaptive Language Learning Quiz for Newcomer 6th Graders. *ACM Trans. Graph.* 37, 4, Article 111 (August 2024), 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Economic and political instabilities worldwide have prompted many families to seek safer environments and better opportunities in new countries. Often, these individuals come from non-English-speaking backgrounds, facing significant linguistic and cultural adjustments. For children, this challenge is particularly acute, as they must learn to thrive in unfamiliar educational settings while overcoming language barriers.

As a 6th-grade bilingual teacher, I have seen firsthand the wide range of linguistic needs among my students. My classroom includes native English speakers, English language learners, and newcomers with limited English proficiency. Addressing this diversity and supporting each student's unique learning trajectory is a persistent challenge.

Educational research highlights the importance of differentiated instruction and personalized learning experiences to foster student success (Tomlinson, 2014). In response, I developed an adaptive language learning quiz designed to enhance the vocabulary and grammar skills of newcomer 6th graders. By providing immediate feedback and adjusting to each learner's current proficiency level, this tool aims to improve the learning experience and ease academic integration.

This project serves as both an educational prototype and a foundation for more comprehensive language learning tools. By experimenting with this initial version, we explore how adaptive technology can support English language acquisition in immigrant student populations, paving the way for future improvements and broader applications.

Following a review of relevant literature and educational theories that underpin adaptive learning, we detail our methodology, including dataset preparation, task generation, and real-time adaptation mechanisms. We then present baseline experiments and results, comparing non-adaptive and adaptive approaches. Finally, we discuss findings, challenges encountered, and potential research directions. Throughout the project, we paid careful attention to aligning our tool with the Texas Essential Knowledge and Skills (TEKS) standards for 6th grade, ensuring its educational relevance.

2 Background and Related Work

Vygotsky's (1978) sociocultural theory introduced the Zone of Proximal Development (ZPD), proposing that learners operate within three conceptual zones. The first encompasses tasks they can perform independently, relying on existing skills. The second, the ZPD itself, includes tasks they cannot yet perform alone but can accomplish with guidance. The third comprises tasks currently beyond their reach, even with help. According to Vygotsky, effective instruction targets the ZPD, enabling learners to advance from reliance on external support to independent mastery.

Scaffolding, an instructional technique stemming from Vygotsky's theory, provides temporary, targeted support gradually withdrawn as learners gain competence (Wood, Bruner, & Ross, 1976). By connecting new material to prior knowledge and using aids like sentence stems or word banks, educators foster increased independence and deeper understanding over time.

The ZPD suggests that learning flourishes when challenges are slightly beyond a learner's current ability, creating a balance between accessibility and stimulation. Adaptive learning technology aligns with this principle, adjusting difficulty and content based on ongoing performance. Immediate feedback in adaptive systems informs the selection of subsequent tasks, ensuring learners continuously engage with material that is neither too easy nor too difficult. Popular platforms like Duolingo, I-Ready, and Khan Academy exemplify such adaptive methodologies (Loewen et al., 2019; Curriculum Associates, 2021; Khan Academy, 2023).

Recent studies indicate that personalizing instruction to a learner's trajectory improves language proficiency and other outcomes (Gerber, 2014; Woolf, 2010). By integrating adaptive learning principles

Author's Contact Information: Jorge Sandoval, UC Berkeley, Dallas, United States of America, sandoval@berkeley.edu.

ACM 1557-7368/2024/8-ART111
https://doi.org/XXXXXXX.XXXXXXX

with Vygotsky's ZPD and scaffolding models, educators can accurately identify a student's readiness level, deliver optimally challenging tasks, and guide learners into their next zone of proximal development.

3 Methodology

This section outlines our data sources, preprocessing steps, and the techniques used to generate, adapt, and deliver language learning tasks. We first discuss dataset preparation, followed by task generation methods. We then describe the adaptive mechanism for real-time difficulty adjustments, and conclude with an overview of the system architecture, environment, and tools.

3.1 Dataset Preparation

Our primary data source was the CommonLit Ease of Readability (CLEAR) Corpus, featuring approximately 5,000 text excerpts graded from 3rd to 12th grade, enriched with metadata such as publication year, genre, author, and various readability metrics. This corpus was chosen due to its alignment with 6th-grade literacy levels and the availability of authentic, varied texts.

Initial exploratory data analysis (EDA) showed 4,724 rows and 28 columns, including numeric, integer, and categorical features. Key readability metrics (e.g., Flesch-Reading-Ease ranging 44.77–81.70 and Flesch-Kincaid Grade Level) helped identify excerpts suitable for 6th-grade learners. Some columns, such as Anthology, Sub Category, and License, had significant missing data and were omitted.

Focusing on text-based tasks, we relied on well-populated columns (Excerpt, Title, readability metrics) and removed less relevant or incomplete fields. We cleaned and normalized the excerpts using Python libraries (pandas, regex) and the Hugging Face datasets library for loading and manipulation. We also employed evaluate and rouge$_s$*corefor performanceassessment.Thefinalpreprocesseddatasetformedastablebasisformodeltrainingandtaskgeneration.*

3.2 Task Generation

With the refined dataset, we developed three primary task types to promote vocabulary, grammar, and contextual understanding.

Baseline (BERT-based) Approach: Initially, we utilized a BERT-based masked language modeling approach (Devlin et al., 2019). Target words were replaced with [MASK], and a fill-mask pipeline predicted likely tokens. This generated plausible distractors for fill-in-the-blank items and identified synonyms. While functional, this baseline lacked contextual nuance and was not tuned for the complexity of 6th-grade material.

Transition to T5: To improve contextual quality, we fine-tuned a T5 model on the CLEAR Corpus. By prompting T5 with specially designed inputs and targets, we achieved more coherent, contextually accurate completions. Using the Hugging Face Transformers ecosystem, tokenization (T5Tokenizer), and text generation (T5ForConditionalGeneration), we integrated complex operations like synonym generation, sentence completion, and distractor filtering aligned with our target proficiency level.

Task Types:

- *Fill-in-the-Blank:* Students complete sentences with missing words chosen from multiple options, reinforcing vocabulary and context understanding.
- *Synonym Matching:* Students select the best synonym for a given word, refining their mental lexicon and comprehension.
- *Sentence Rearrangement:* Students reorder scrambled sentences, enhancing discourse coherence and logical reasoning skills.

Shifting from a BERT-based baseline to a fine-tuned T5 model enhanced educational robustness, more closely mirroring authentic language complexity and supporting adaptive learning.

3.3 Adaptive Mechanism

The adaptive mechanism adjusted difficulty based on student performance and response times:

- *Accuracy:* Consistent correctness led to incremental difficulty increases, with more advanced vocabulary and subtler distractors.
- *Response Time:* Extended hesitation suggested difficulty, prompting simplification via clearer options or lower-complexity passages.

Feedback System: Immediate feedback informed learners if their answers were correct, providing explanations, synonyms, or clarifications. This feedback loop encouraged reflection, reinforcing proper usage and guiding learners toward improved accuracy, ultimately scaffolding them through their ZPD.

3.4 System Architecture

The system integrated data preprocessing, T5-based task generation, and performance tracking into a unified workflow. Implemented in Python and leveraging Google Colab's GPU capabilities, it utilized Hugging Face's Hub for pretrained models, pandas, matplotlib, seaborn for EDA, spacy and nltk for NLP, and evaluate and rouge$_s$*corefor assessment*.

4 Experiments and Results

4.1 Baseline System

The BERT-based baseline achieved around 60% accuracy but lacked adaptivity. Although it generated fill-in-the-blank items and distractors, its static difficulty yielded no long-term engagement or performance gains. It provided a proof of concept but failed to evolve with student proficiency, limiting its educational impact.

4.2 Adaptive System Evaluation

Evaluating the adaptive T5-based approach on a held-out subset, we used accuracy, BLEU, and ROUGE metrics. Compared to the baseline, the adaptive system scored higher across these measures, indicating more contextually and semantically appropriate outputs. By adjusting difficulty and providing immediate feedback, it offered a more responsive and effective learning experience.

4.3 Adaptive System Success and Challenges

As students answered correctly and swiftly, the system increased difficulty, attempting to approach each learner's ZPD. However, complexity sometimes rose too quickly, challenging even native speakers. This indicates a need for more refined calibration. Additionally, error analysis revealed issues like imprecise synonyms,

overly simple targets, ambiguous choices, and tasks with multiple plausible solutions.

5 Discussion

While the adaptive T5-based quiz surpassed the static baseline, its results were mixed. Enhanced accuracy, BLEU, and ROUGE confirmed richer contextual tasks, yet it did not fully embody Vygotsky's ZPD or implement robust scaffolding. Some outputs lacked nuanced appropriateness, and not all tasks matched evolving learner proficiency.

These limitations may stem from insufficient feedback loops, lack of in-classroom testing, and incomplete integration of scaffolding strategies. Without longitudinal data, long-term efficacy remains unclear. Nonetheless, this prototype offers insight into improving adaptive strategies, refining adaptation rules, and incorporating richer hints or prompts to achieve more meaningful language practice.

6 Conclusion

We developed an adaptive, T5-based language quiz for newcomer 6th graders, demonstrating promise yet acknowledging the need for refinement. While not consistently delivering "just right" challenges, it provides a valuable foundation.

6.1 Summary of Contributions

We introduced an adaptive language quiz aligned with TEKS, surpassing static approaches by adjusting difficulty and offering immediate feedback. This lays groundwork for more personalized, theoretically informed language learning tools.

6.2 Educational Impact

With further refinement, this prototype could integrate seamlessly into classrooms, supporting English learners as they adapt to new academic environments and improving their language acquisition trajectory.

6.3 Future Work

Future research may explore reinforcement learning for subtler adaptations, include listening and speaking tasks, and conduct longitudinal studies to measure genuine proficiency gains. Collaborations with educators and iterative classroom trials will guide ongoing enhancements.

6.4 Concluding Remarks

Though not yet perfected, this prototype illustrates the potential of adaptive technology in language education. With sustained refinement and evaluation, we may approach truly effective, data-driven solutions that enhance language learning experiences for newcomers and beyond.

Acknowledgments

We would like to thank those who supported this project and provided feedback on the development of the adaptive language learning quiz.

References