EUROCALL 2023: CALL for all Languages
15-18 August 2023, University of Iceland, Reykjavik
DOI: https://doi.org/10.4995/EuroCALL2023.2023.16996

The use of machine learning in developing learner-adaptive tools for second language acquisition

Maryam Sadat Mirzaei[a] and Kourosh Meshgi[b]
[a]RIKEN Center for Advanced Intelligent Project, , maryam.mirzaei@riken.jp and [b]RIKEN Center for Advanced Intelligent Project, , kourosh.meshgi@riken.jp

How to cite: Mirzaei, M.S.; Meshgi, K. (2023). The use of machine learning in developing learner-adaptive tools for second language acquisition. In *CALL for all Languages - EUROCALL 2023 Short Papers.* 15-18 August 2023, University of Iceland, Reykjavik. https://doi.org/10.4995/EuroCALL2023.2023.16996

Abstract
Advancements in artificial intelligence and machine learning present opportunities to revolutionize language learning tools with learner-adaptive capabilities. These technologies facilitate the creation of trainable systems that can interact with learners, offering personalized learning experiences tailored to individual needs, interests, proficiency levels, backgrounds, and native languages. This study explores the role of machine learning in developing personalized frameworks for second language learning, introducing the Partial and Synchronized Caption (PSC) tool as an example. PSC utilizes automatic speech recognition and natural language processing to identify challenging words for language learners, which are presented in the caption while masking easy words. We used machine learning to personalize the caption for various learners. An experiment involving graduate students learning English as a second language demonstrated the adaptability of PSC's word selection to different learners. While creating entirely personalized captions may be challenging, PSC offers a promising approach to personalized and adaptable listening tools. The data collected from learner interactions also provides valuable insights into individual needs, shaping future language learning tools and pedagogical practices.

Keywords: *Machine learning, learner-adaptive technologies, personalized language learning, Partial and Synchronized Caption.*

1. Introduction

Machine learning has enabled the production of personalized technologies that focus on individual users' preferences and needs, emerging in the form of recommendation systems, personal AI assistants, news feeds, and many more. Recently, leveraging data-driven algorithms has enabled tailoring learning experiences by providing personalized, custom-made learning practices for individual learners based on their unique needs, interests, and proficiency levels, especially in the domain of second language learning (Chen et al., 2021). The main goal is to deliver: (a) precise and appropriate content; (b) accurate assessments; (c) tailored feedback; and (d) a customized learning path. In this context, many studies have focused on the development of learner-centric tools and content personalization to address diverse learners' preferences. These studies cover a range of applications, including personalized readability assessment for L2 reading (Ehara, 2022), difficulty detection and adaptation for practicing grammar (Pandarova et al., 2019), personalized conversational AI agents (Dizon et al., 2022), robot-assisted language learning (Randall, 2019), and simulation and games (Karoui et al., 2021; Peterson & Jabbari, 2022). Notably, there is a specific focus on personalized mobile-assisted learning (Gumbheer et al., 2022).

The development of personalized profiles or learner models facilitates effective content filtering, directing learners toward resources that match their specific proficiency levels, learning objectives, and content preferences (Godwin-Jones, 2017). Personalization occurs at different levels. Traditionally adaptive systems focused on coarse-grained personalization, providing general and broad-level personalization by drawing on user demographics and grouping users into broader categories based on shared characteristics, preferences, or behaviors (Walkington & Bernacki, 2020). The advantages include easier implementation and reduced data requirements within these broad categories. However, more complex adaptation is needed to provide accurate and tailored learning practice that is aligned with individual learners' needs (Ismail et al., 2016). The fine-grained personalized system addresses these limitations by considering the unique preferences and behavior of each user, accounting for specific interactions, historical data, and individual feedback to provide a highly personalized learning experience. Nonetheless, incorporating learner diversity requires more individual user data. Furthermore, personalization can be combined with in situ learning to provide personalized content that is contextually and situationally relevant (Dalton-Puffer & Smit, 2013). Finally, dynamic and progressive personalization includes the continuous update of the system based on the learner's progress and evolving needs for long-term practice and learning outcomes (Shute & Rahimi, 2017).

To provide an example, we introduce a tool for L2 listening development called Partial and Synchronized Caption (PSC). With this method, we aim to move toward a learner-centered listening practice, with the ability to fine-tune the system based on the learner's preferences and advancement constantly.

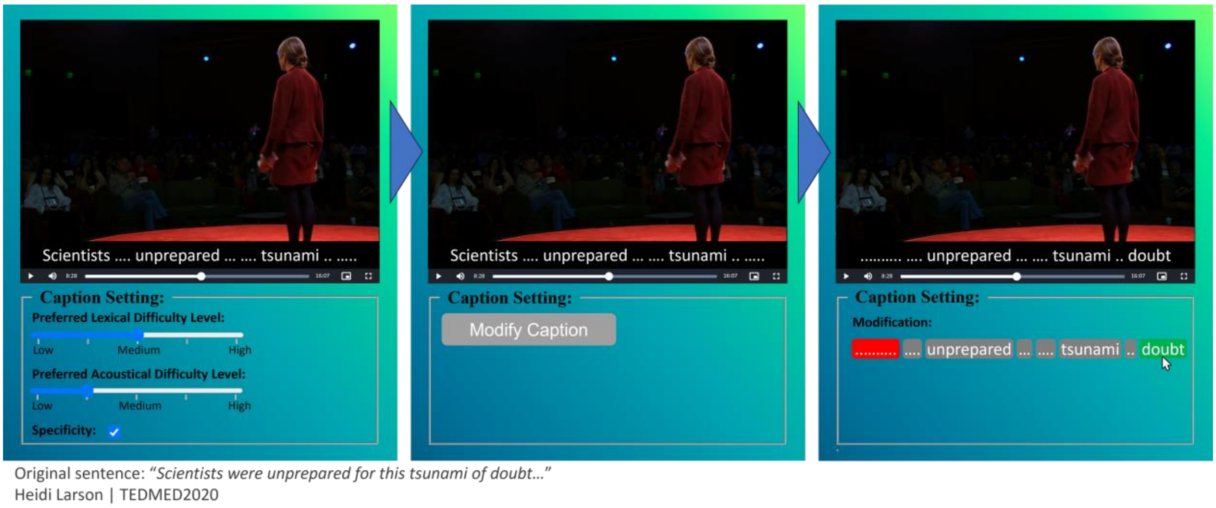

Figure 1. PSC self-regulation feature (left) and word choice personalization feature (middle and right)

2. Personalized caption

PSC uses lexical and acoustic features to determine the difficulty level of the words and decides which parts of the text need to be presented in the caption, intending to encourage listening over reading. The caption is synchronized on the word level, thereby enabling seamless audio and text mapping and reducing distraction.

For the baseline version, we used rule-based coarse-grained level assignments to roughly categorize learners into three language proficiency levels (beginners, intermediate, advanced) based on learner's assessment tests (TOEFL/TOEIC score, speech rate tolerance, vocabulary size). Word selection is determined by defining thresholds for specific features, including word frequency and speech rate, while also incorporating additional factors like automatic speech recognition system errors, word specificity, proper names, and abbreviations (Mirzaei et al., 2018). Compared to the full caption and keyword caption, the baseline version of PSC provides a certain level of personalization by adapting the words in the caption to different proficiency levels. Yet, within each group, learner variability poses a challenge, requiring further adaptation to meet individual user preferences and requirements. Therefore, we employed two distinct approaches that integrate two modes of learner feedback (Figure 1).

2.1. Self-paced learning through learner customization

Self-directed learning is crucial, emphasizing individuals' initiative in diagnosing needs, adapting strategies, and evaluating outcomes. For listening, self-paced learning personalizes the experience, fostering agency, and motivation, and reducing cognitive overload, leading to improved performance (Ozcelik et al., 2019). Acknowledging that learners may have varying strengths in vocabulary retention and tolerance for fast speech, we have integrated a user-friendly interface, where users can easily modify the system's parameters and adjust the number of words displayed, hence promoting autonomous and self-regulated learning. This level of customization allows learners to tailor the language learning process to their unique preferences, pace, and proficiency. Those with a robust vocabulary can opt for a fewer number of words to challenge themselves further. On the other hand, individuals who struggle with fast speech can increase the number of words with a faster speech rate displayed, allowing them to focus on mastering the content at a pace that suits their comfort level. The flexibility of the learning environment encourages learners to explore and experiment freely.

2.2. Personalization through learner feedback

This subsequent stage involves automatic personalization or adaptation. During this phase, we gather learners' feedback to enhance the word choices within the caption and align it with their preferences. If a learner opts to modify the word selection, the words within the caption become visible in clickable boxes. At this point, if a learner finds a particular word unnecessary in the caption, that word hides and is marked for subsequent system retraining. Conversely, if a learner wishes to reveal a hidden word, a simple click unveils it, prompting the system to update accordingly.

To achieve adaptation, the system computes the differences between the original generated caption and any modifications introduced to it. Following this, the system adjusts its parameters to optimally align with the marked words by dynamically fine-tuning the coefficients of the extracted features, utilizing supervised learning. This iterative process ensures that the system continuously improves its personalized word selection based on user interactions and feedback. The classifier is retrained with this data, allowing it to grasp insights into each individual's language learning challenges, background, vocabulary knowledge, and potential sources of listening difficulties to enhance personalized captions.

3. Experiments

We conducted a preliminary experiment with this tool involving 29 participants who were intermediate English learners. They were graduate students, aged 22-26, from diverse academic backgrounds including fields such as medicine, business, and engineering. The material used for the experiment consisted of TED talks delivered by native English speakers.

Participants accessed the experimental session online. Ahead of the experiment, they acquainted themselves with the system by viewing three brief videos and testing the self-regulation (adjusting default parameters) and feedback collection (editing shown/hidden words in the caption) functionalities. After this introductory episode, the first stage of the experiment involved watching four videos (V1-V4) with baseline PSC, where participants could customize the system parameters for self-regulated listening. They adjusted the frequency and speech rate threshold for the generated caption during the first five minutes of each video, creating their desired settings. Whenever they faced listening difficulty, they could pause the video and use the clickable box to see the hidden words they couldn't recognize. They could also omit unnecessary/distracting words. The system stored this user feedback to fine-tune its future word selection and generate captions that aligned more with user preferences. After each video, participants received a report detailing the percentage of shown/hidden words in their personalized caption, motivating them to remove unwanted words to provide room for desired ones.

In the second stage of the experiment, the system used learner data from the previous videos to better adapt its word selection for the subsequent four videos (V5-V8). Participants continued to use the modification function to show/hide words when necessary. The learners' log files from both stages were stored for further analysis. We also had a brief interview with learners to capture any feedback or suggestions.

4. Results and discussion

Figure 2 illustrates learner feedback on the system, analyzed through log files across the two stages of the experiment, using distinct videos. In the initial stage, participants actively engaged with the system, refining captions to align with their proficiency and demands. The feedback collection feature allowed for pinpointing challenging words and eliminating unnecessary ones, resulting in more personalized captions. Notably, interaction peaked at this point, suggesting active customization of captions to match learners' expectations. The focus was primarily on shown words, likely due to their impact on listening difficulties. However, as depicted in the figure, the volume of learner modifications decreased as the experiment progressed to the next stage.

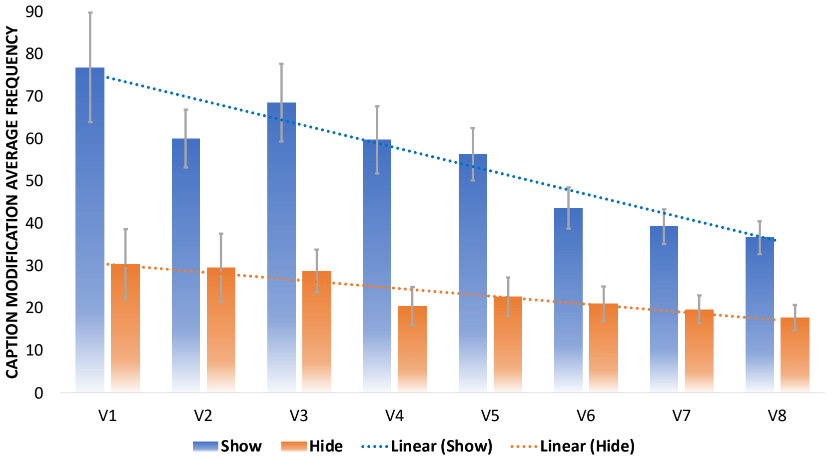

Figure 2. Experimental result on learners' modification of PSC's word choices

The second stage showcased the PSC system's adaptability (V5-V8). Leveraging data from the initial stage, the system further improved its word selection for subsequent videos, enhancing overall personalization. Fewer modifications in this stage underscores the PSC's ability to adapt its word selection, capturing essential words for individual learners to be shown in the caption. Nevertheless, there remained potential for improvement by omitting unnecessary words based on learner feedback.

We discuss two primary factors in Figure 2. One pertains to **learner variability** while the other is associated with **content variability**. The experiment's outcome highlights differences in show/hide word feedback across our learners. The observed deviation (STD bar in the figure) in the number of modifications made by learners highlights learner variability, illustrating that some needed more adjustments while others required fewer. This emphasizes the diversity among learners within the same group.

When examining the videos themselves, it becomes apparent that, for certain videos, modifications increased (V3&V5), and most learners required an increased number of shown words. This observation aligns with the learners' feedback during the interview phase where most learners indicated that these videos were notably challenging, featuring either fast speech rate or technical language, and subject complexity. However, two learners did not express particular difficulty with these videos as they shared a background related to the video's topic. This finding is important, as it introduces another factor (learner background) that should be considered in the context of adaptation.

The interview data also indicated that certain learners utilized the modification feature to reveal hidden words, to confirm their correct recognition, after which they opted to hide the word again. Conversely, this also implies that learners might hide some words under the assumption of their ease of recognition, potentially influenced by having seen those words in the caption. This suggests that learners might not always have a comprehensive awareness of their requirements for self-regulation. Additionally, learner feedback provided valuable insights into interface improvements, the inclusion of repetition, and the ability to select preferred videos.

5. Conclusions

While creating a caption that is truly designed for each learner may not be entirely feasible, we found that PSC could adapt its word selection to an acceptable level for different learners. This presents a promising avenue for crafting personalized, adaptable listening tools using machine learning, automatic speech recognition, and natural language processing. Notably, learner and content variations significantly impact personalization and system efficacy. It's worth noting that if adaptable technology falls short of user expectations, it can detrimentally affect satisfaction and engagement. Finally, our results underscore the need to expand findings across various proficiency levels and backgrounds as well as compare with other existing methods, such as keyword captioning in future research. Additionally, tailoring feedback and enabling reflection and self-monitoring are essential considerations.

References

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning. *Educational Technology & Society*, *24*(1), 205-222.

Dalton-Puffer, C., & Smit, U. (2013). Content and language integrated learning: a research agenda. *Language Teaching*, *46*(4), 545-559. https://doi.org/10.1017/s0261444813000256

Dizon, G., Tang, D., & Yamamoto, Y. (2022). A case study of using Alexa for out-of-class, self-directed Japanese language learning. *Computers and Education: Artificial Intelligence*, *3*, 100088. https://doi.org/10.1016/j.caeai.2022.100088

Ehara, Y. (2022). Uncertainty-aware Personalized Readability Assessment Framework for second language learners. *Journal of Information Processing*, *30*, 352-360. https://doi.org/10.2197/ipsjjip.30.352

Godwin-Jones, R. (2017). Scaling up and zooming in: Big data and personalization in language learning. *Language Learning & Technology*, *21*(1), 4–15. https://dx.doi.org/10125/44592

Gumbheer, C. P., Khedo, K. K., & Bungaleea, A. (2022). Personalized and adaptive context-aware mobile learning: review, challenges and future directions. *Education and Information Technologies*, *27*(6), 7491-7517. https://doi.org/10.1007/s10639-022-10942-8

Ismail, H. M., Harous, S., & Belkhouche, B. (2016). Review of personalized language learning systems. In *2016 12th International Conference on Innovations in Information Technology* (pp. 1-6). IEEE. https://doi.org/10.1109/innovations.2016.7880051

Karoui, A., Alvarez, L., Goffre, T., Dherbey Chapuis, N., Rodi, M., & Ramalho, M. (2021). Adaptive pathways within the European platform for personalized language learning PEAPL. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 90-94). https://doi.org/10.1145/3450614.3464480

Mirzaei, M. S., Meshgi, K., & Kawahara, T. (2018). Exploiting automatic speech recognition errors to enhance partial and synchronized caption for facilitating second language listening. *Computer Speech & Language*, *49*, 17-36. https://doi.org/10.1016/j.csl.2017.11.001

Ozcelik, H. N., Van den Branden, K., & Van Steendam, E. (2019). Listening comprehension problems of FL learners in a peer interactive, self-regulated listening task. *International Journal of Listening*, 1-14. https://doi.org/10.1080/10904018.2019.1659141

Pandarova, I., Schmidt, T., Hartig, J., Boubekki, A., Jones, R. D., & Brefeld, U. (2019). Predicting the difficulty of exercise items for dynamic difficulty adaptation in adaptive language tutoring. *International Journal of Artificial Intelligence in Education, 29*, 342-367. https://doi.org/10.1007/s40593-019-00180-4

Peterson, M., & Jabbari, N., eds. (2022). *Digital games in language learning: case studies and applications.* Taylor & Francis. https://doi.org/10.4324/9781003240075-1

Randall, N. (2019). A survey of robot-assisted language learning (RALL). *ACM Transactions on Human-Robot Interaction (THRI), 9*(1), 1-36. https://doi.org/10.1145/3345506

Shute, V. J., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning, 33*(1), 1-19. https://doi.org/10.1111/jcal.12172

Walkington, C., & Bernacki, M. L. (2020). Appraising research on personalized learning: definitions, theoretical alignment, advancements, and future directions. *Journal of research on technology in education, 52*(3), 235-252. https://doi.org/10.1080/15391523.2020.1747757