# AI Misalignment in Language Learning: Investigating Conflicts Between Automated Feedback and Pedagogical Intentions

**Irina Vardanyan**
*Yerevan State University, Armenia*

## Abstract

With the rapid growth of AI in foreign language teaching, automated feedback systems are considered an invaluable means of providing immediate and individualised feedback. However, little attention has been paid to the situations where AI-generated feedback diverges from the teacher's pedagogical intentions. This study examines the phenomenon of pedagogical misalignment, where accuracy-driven automated corrections hinder teachers from developing communicative fluency, interactional competence, or task-based language performance. Drawing on a mixed set of classroom data collected over six weeks, comprising lesson observations, teacher interviews, and screen-capture recordings from pre-intermediate and intermediate English classes, the study reveals recurring patterns of misalignment that disrupted instructional flow, produced learner uncertainty, and occasionally challenged teachers' authority. Findings reveal that AI tools frequently emphasised minor errors, suggested corrections which didn't correspond with the task objectives and interrupted learners during fluency-oriented activities. The article outlines how teachers responded to these tensions, including clarifying AI feedback for learners, modifying task procedures, and explicitly guiding students in interpreting automated corrections. This study presents a pedagogical alignment model specifying when AI-generated feedback should be applied, adapted through teacher mediation, or intentionally limited. Overall, the results demonstrate that purposeful and context-sensitive integration of AI feedback supports more coherent communicative pedagogy while protecting learner agency in AI-enhanced classrooms.

**Keywords:** AI feedback; TESOL; pedagogical alignment; communicative competence; classroom interaction