Why generative AI cannot erase grammar: Persistent L1 effects in L2 tense-aspect under generative AI editing

Yuichi Ono* and Yuka Fujii

University of Tsukuba, ono.yuichi.ga@u.tsukuba.ac.jp*

Keywords: learner corpus research; generative AI; native language identification; tense-aspect; Feature Reassembly Hypothesis

This study investigates whether generative AI editing can mitigate first-language (L1) effects on tense-aspect usage in L2 English, using cross-linguistic typology and the Feature Reassembly Hypothesis (FRH) as a theoretical framework (Lardiere, 2008, 2009; Slabakova, 2008). Previous research has shown that acquisition of tense and aspect is strongly constrained by how these meanings are grammaticized in the learner's L1 (Bardovi-Harlig, 2000; Shirai, 2024), with Japanese employing the *-te iru* construction, which is ambiguous between progressive and resultative readings (Sugaya & Shirai, 2007), and Chinese lacking an explicit grammatical tense marker, while Korean provides more transparent counterparts for past, progressive, and resultative/perfect-like constructions. Feature Reassembly predicts that features without clear L1 correspondents—most notably the English perfect—should be especially resistant to restructuring (Lardiere, 2008; Slabakova, 2008).

We use generative AI editing as an external intervention to probe the stability of interlanguage representations. Argumentative essays produced by Chinese (CHN), Japanese (JPN), and Korean (KOR) learners of English, together with native speakers (NS), were drawn from the *International Corpus Network of Asian Learners of English* (ICNALE; Ishikawa, 2023; n = 2,312 documents). Each essay underwent three levels of automated editing (*light, normal, strong*), ranging from minimal grammatical correction to fluency-oriented revision and native-like sentence restructuring, including tense-aspect adjustment. Binary presence/absence indicators for *past, present, perfect*, and *progressive* were extracted for each version, and AI-induced grammatical change was modeled at both the document level (any tense-aspect change) and the category level (past, perfect, progressive) using generalized linear mixed-effects models.

Across all editing stages, AI modified 2,562 past, 2,014 perfect, and 2,585 progressive constructions. Of these, 970 past (37.9%), 481 perfect (23.9%), and 255 progressive tokens (9.9%) were eliminated in the strong editing condition, indicating qualitatively different patterns of category-level modifiability. Progressive constructions were rarely deleted and instead tended to be redistributed, whereas past showed mixed deletion and restructuring. By contrast, perfect constructions were both less frequently targeted overall and more likely to be avoided rather than reassembled, consistent with their categorical resistance observed in the mixed-effects models.

Results demonstrate statistically robust persistence of L1-conditioned patterns despite increasingly aggressive AI rewriting. At the document level, KOR consistently aligned with NS, whereas CHN and JPN showed markedly lower probabilities of change across stages (e.g., light→normal: KOR vs. JPN, OR = 4.24, 95% CI [3.42, 5.26]). Category-level analyses revealed a resistance hierarchy (perfect > past > progressive).

These findings align closely with Feature Reassembly predictions (Lardiere, 2008, 2009; Slabakova, 2008): forms lacking transparent L1 mappings—most notably the English perfect for Japanese and Chinese—were most resistant, whereas Korean learners patterned natively, consistent with earlier work on near-native grammatical competence and aspectual contrasts (Montrul & Slabakova, 2003). Crucially, robust L1 effects persisted even under strong editing that permitted sentence reorganization and tense–aspect adjustment, indicating that generative AI primarily operates on surface distributions rather than inducing deep grammatical reassembly. From an LCR perspective, these findings suggest how combining learner corpora with controlled AI editing provides a scalable method for probing interlanguage stability beyond traditional experimental paradigms.

References

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Wiley-Blackwell.

Ishikawa, S. (2023). *The ICNALE guide: An introduction to learner corpus study on Asian learners' L2 English*. Routledge.

Lardiere, D. (2008). Feature assembly in second language acquisition. In J. M. Liceras, H. Zobl, & H. Goodluck (Eds.), *The role of formal features in second language acquisition* (pp. 106–140). Lawrence Erlbaum Associates.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research, 25*(2), 173–227.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in Second Language Acquisition, 25*(3), 351–398. https://doi.org/10.1017/S0272263103000159

Shirai, Y. (2024). *Tense and temporality*. In T. Ionin, S. Montrul, & R. Slabakova (Eds.), The Routledge handbook of second language acquisition, morphosyntax, and semantics (pp. 257–269). Routledge.

Slabakova, R. (2008). *Meaning in the second language*. De Gruyter Mouton.

Sugaya, N., & Shirai, Y. (2007). The acquisition of progressive and resultative meanings of the imperfective aspect marker by L2 learners of Japanese: Transfer, universals, or multiple factors? *Studies in Second Language Acquisition, 29*(1), 1–38. http://www.jstor.org/stable/44488643