Lexical Event Structure and the Distribution of L2 Tense and Aspect Errors

Carlos Balhana
Department of Theoretical and Applied Linguistics
University of Cambridge

Over nearly 30 years, sustained research activity on the Aspect Hypothesis (Andersen & Shirai, 1994) has firmly coalesced around the position that the emergence of L2 tense and aspect marking is conditioned by lexical aspect early in learners' trajectories (Bardovi-Harlig & Comajoan-Colomé, 2020). For L2 English, the hypothesis maintains that learners will initially mark past tense correctly on achievement and accomplishment verbs, which describe events with instantaneous or foreseeable endpoints (e.g. *break* [*a vase*], *build* [*a house*], *bake* [*a cake*], *find* [*your keys*], etc.), before doing so reliably with activity and state verbs, whose event depictions have no predetermined or prefigured conclusion (e.g. *walk*, *sleep*, *love*, *study, know,* etc.) (Bardovi-Harlig, 1999). The hypothesis further predicts that learners will mark progressive aspect on activities before achievements and accomplishments and will refrain from ungrammatically marking it on states (#*I am knowing* [*that…*], etc.). In this talk, we will explore a slightly different take on the influence of lexical aspect and examine whether event structure might affect the production of tense and aspect targets. Although the Aspect Hypothesis does not commit itself to predicting morphological errors on the basis of verb class (Shirai, 2007), we might posit that lexical aspect can not only prime but also interfere with form-meaning mappings of tense and grammatical aspect; if so, we would expect to find overproduction of past and perfect aspect surface errors with achievements and accomplishments and overproduction of present and progressive surface errors with activities and states.

For our analysis, we first collected over 900,000 L2 learner task-based writing samples from the EFCAMDAT corpus (Geertzen et al., 2013), spanning 7 different L1 backgrounds across basic and intermediate proficiency levels (~A1-B1). We then applied natural language processing (NLP) techniques to automatically label tense and aspect features for both the student-produced verb forms and, in the case of errors, instructor-provided targets. Assuming an event-based approach to verb meaning, we then enriched these text annotations with predicate-level semantic representations. Under a truly compositional view, achievements and accomplishments may subsume atelic predicates in their event structure – the semantic representation of the (achievement) verb *sell*, for example, includes the stative relation HAS_POSSESSION. Finally, we performed a series of statistical tests and modeling studies to evaluate if L2 learners are sensitive to these structures and how they might commingle lexical and functional layers of the interlanguage during morphosyntactic realization. Our findings suggest that event representations are strong predictors of tense and aspect errors and we propose some future applications for online language learning.

Andersen, R. W., & Shirai, Y. (1994). Discourse Motivations for Some Cognitive Acquisition Principles. *Studies in Second Language Acquisition*, *16*(2), 133–156. https://doi.org/10.1017/S0272263100012845

Bardovi-Harlig, K. (1999). From Morpheme Studies to Temporal Semantics: Tense-Aspect Research in SLA. *Studies in Second Language Acquisition*, *21*(3), 341–382.

Bardovi-Harlig, K., & Comajoan-Colomé, L. (2020). The aspect hypothesis and the acquisition of L2 past morphology in the last 20 years: A state-of-the-scholarship review. *Studies in Second Language Acquisition*, *42*(5), 1137–1167.

Geertzen, J., Alexopoulou, T., & Korhonen, A. (2013). Automatic linguistic annotation of large scale L2 databases: The EF-Cambridge Open Language Database (EFCAMDAT). *Proceedings of the 31st Second Language Research Forum. Somerville, MA: Cascadilla Proceedings Project*, 240–254.

Shirai, Y. (2007). The Aspect Hypothesis, the comparative fallacy and the validity of obligatory context analysis: A reply to Lardiere, 2003. *Second Language Research*, *23*(1), 51–64.