

Expanding Learning Environments with Data from Learner-computer Interactions

¹Trude Heift, ²Catherine Caws

¹Simon Fraser University, ²University of Victoria (Canada)

heift@sfu.ca , ccaws@uvic.ca

Abstract

Data collections from learner-computer interactions are commonly used to determine learning outcomes and/or improve software features that are designed to enhance the learning experience for specific learning activities. For this reason, a cyclical approach to software engineering [Colpaert, 2006, Hubbard, 2011] is generally preferred because each and every stage during the lifecycle delivers output that serves as input for the subsequent stage. Accordingly, these data collections and analyses of learner data have provided recommendations on the efficacy of certain learning tools as well as their technical usability by also highlighting the needs for user testing and feedback for system improvement. A far less explored application, however, has emphasized the recycling of learner data to expand an existing learning environment (e.g., Tutorial CALL) with data-driven learning (DDL), especially if the corpus has been constructed from authentic learner submissions of the very same learning environment.

DDL has been explored under various aspects by generally emphasizing the ways in which it can facilitate the implementation of a methodology for language learning that focuses on authenticity in contents, context, and task [van Lier, 1996], commonly achieved with a corpus of L1 or L2 data. This presentation focuses on DDL built around the recycled learner data of E-Tutor, a comprehensive ICALL environment for German.

Over a time period of five years, millions of user submissions from a variety of learning tasks in E-Tutor were collected and from those a common learner corpus was constructed. Due to the underlying NLP analysis and tagging capabilities of E-Tutor, the learner corpus has multiple applications. Firstly, learners can examine interlanguage or task-specific phenomena and the benefits in this respect have been well documented [Braun, 2005; Granger, 2003]. Secondly, learners can compare their concurrent interlanguage data with that of previous users and determine the ways in which their interlanguage is alike or differs. Thirdly, it allows language instructors and/or researchers to examine the design of language learning material given that the corpus reveals much about the degree of difficulty of each task. Finally, these large data collections allow for investigations of theoretical research topics (e.g., interlanguage and task-related studies) to be explored by the researcher. This presentation will discuss the multiple applications of DDL and user studies that consider these corpus explorations by learners, teachers, and researchers alike.

References

- [1] Braun, S. (2005). From Pedagogically Relevant Corpora to Authentic Language Learning Contents. *ReCALL*, 17(1), 47-64.
- [2] Colpaert, J. (2006). Pedagogy-driven design for online language teaching and learning. *CALICO Journal*, 23(3), 477.
- [3] Granger, S. (2003). Error-tagged Learner Corpora and CALL: A Promising Synergy. *CALICO Journal*, 20(3), 465-480.
- [4] Hubbard, P. 2011. "Evaluation of Courseware and Websites." In L. Ducate and N. Arnold (Eds.) *Present and Future Perspectives of CALL: From Theory and Research to New Directions in Foreign Language Teaching*, Second Edition. San Marcos, TX: CALICO.
- [5] Van Lier, L. (1996). *Interaction in the Language Curriculum: Awareness, Autonomy, and Authenticity*. London: Longman.