Error Diagnosis and Error Correction in CALL

Trude Heift
Simon Fraser University

Mathias Schulze
University of Waterloo

This special issue of the CALICO Journal concentrates on an important aspect of learner language—linguistic errors. Recently, issues like error description and classification as well as error correction and feedback have received increased attention in a number of academic disciplines, for example, second language acquisition theories and Corpus Linguistics. Recent conferences in CALL have also seen a number of papers which address these problems. This issue of the CALICO Journal aims to provide an overview of current relevant research. It contains nine articles that deal with a variety of questions focusing on one common theme: learner errors in CALL.

It is noteworthy that the majority of CALL systems described here implement natural language processing (NLP) with the common goal of (a) processing learner errors more effectively and/or (b) developing CALL materials that focus on learners' attested difficulties. The advent of the world wide web and, to some extent, of CD-ROMs has resulted in an overwhelming amount of technologically innovative, multimedia CALL materials. At the same time, it appears that this development has also led to a reduced emphasis on pedagogical issues, one of which is learner feedback. We hope that this special issue is evidence that there is not only interest in applying and developing CALL materials that provide meaningful feedback to the learner but that there are also some promising results made with NLP.

Our special issue begins with an article by Schulze who discusses learner feedback from a theoretical point of view. In order to draw conclusions for CALL system design, he considers terms such as feedback, reinforcement, dialog and their entailments in three disciplines: learning psychology, second language acquisition theory, and human-computer interaction. Examples are taken from his parser-based CALL system Textana for English-speaking learners of German.

Cowan, Choi, and Kim address four questions relevant to error diagnosis and correction in CALL. The questions relate to the diagnosis of persistent L2 learner grammar errors: whether they can be corrected, what types of feedback from the computer are most efficient for focusing the students' attention on this
Error Diagnosis and Error Correction in CALL

task, and the assessment of CALL programs designed to correct errors. A large corpus of L2 learner errors is shown to be highly beneficial for identifying persistent L1 transfer errors. By way of example, the application of concordancing programs for this purpose is demonstrated with Korean learners of English.

Granger also discusses the benefits of learner corpora as a new resource for second language acquisition and foreign language teaching specialists. She describes the FRIDA corpus, an error-tagged corpus for French as a second language that is used to extract detailed error statistics and to carry out concordance-based analyses of specific error types. The results were used to focus CALL exercises on learners’ attested difficulties and to improve the error diagnosis system integrated into the CALL program which is described by L'haire and Vandeventer Faltin. They present an overview of the research conducted within the FreeText project in order to build an automatic error diagnosis system for learners of French as a foreign language. After a brief review of the main features of the project, the paper focuses on the error diagnosis system itself.

A different error recognition module of an interactive ICALL system is discussed by Reuer who focusses on the underlying grammar theory. He argues that the theory of lexical functional grammar (LFG) is well suited to be used in the parsing and error recognition module of the system as well as to provide intelligent feedback to learners. He asserts that the concepts and structures used in LFG closely resemble the descriptive knowledge of language learners about a language.

Delmonte’s systems also rely mainly on LFG. He presents four NLP-based exercise sets where error correction and feedback are produced by means of a rich database. He shows that a parser can be put to use in a variety of pedagogical settings for different languages and for language learners and students of Linguistics alike.

Aspects of another NLP-based system—The German Tutor—are at the center of attention in Heift’s contribution to this volume. She discusses the obstacles of multiple learner errors. Motivated by pedagogical and linguistic design decisions, her web-based ICALL system for German ranks student errors by way of a flexible Error Priority Queue: the grammar constraints can be reordered to reflect the desired emphasis of a particular exercise. In addition, a language instructor might choose not to report some errors. Heift presents a study that supports the need for a CALL system that addresses multiple errors by considering language teaching pedagogy.

Tschichold, who takes a more critical look at NLP in CALL, describes the obstacles parser-based error detection faces in its attempts to provide intelligent feedback in CALL. To overcome some of these obstacles, she proposes a shift in focus to lexical items, both simple words and multiword units of various types. Single and multiword lexemes should not only be explicitly taught in CALL, but could also provide the key to more effective feedback on the language production by learners.
The final paper of this special issue is by Chen and Tokuda, who discuss a new template-template scheme to simplify and streamline the labor-intensive input of possible answers. They argue that their approach, which employs both traditional pattern matching through answer templates and NLP, solves a bottleneck problem of authoring a finite state automaton-based ICALL system capable of automatically correcting free-format English sentences produced by Japanese learners.

ACKNOWLEDGMENTS

We would like to thank the authors for their contribution to this edition who made our work interesting. We are also very grateful to the reviewers—L. Borin, J. Burston, T. Chanier, P. Gupta, B. Harriehausen, V. Hegelheimer, G. Holmes, W. Menzel, N. Nagata, A. Ramsay, A. Sanders, and R. Sanders—who made our task an easier one. Last but not least, we thank the CALICO Editorial Board who made this special issue possible.

EDITORS’ BIODATA

Trude Heift is an Assistant Professor in the Linguistics Department and the Director of the Language Learning Centre at Simon Fraser University. Her research areas are in CALL, Computational Linguistics, and Applied Linguistics. Her main interests are in ICALL, human-computer interaction, feedback, student modeling, and error analysis. Currently, she is working on learner feedback and exploring different help options in ICALL systems.

Mathias Schulze’s research focus is the application of linguistic theory to CALL. He received his Ph.D. in Language Engineering (CALL and German Linguistics) at UMIST in Manchester/England. He has published on language technology in CALL and the acquisition of grammar through CALL. Currently, he is working on the computational implementation of German grammar, computerized adaptive language testing, and the online teaching and learning of German. He is an Assistant Professor of German in the Department of Germanic and Slavic Studies at the University of Waterloo.
EDITORS’ ADDRESSES

Trude Heift
Linguistics Department
Simon Fraser University
Burnaby, B.C.
Canada V5A 1S6
Phone: 604/291-3369
Fax: 604/291-5659
Email: heift@sfu.ca

Mathias Schulze
University of Waterloo
Department of Germanic and Slavic Studies
200 University Avenue West
Waterloo, ON
Canada N2E 3G1
Tel: 519/888-4567, x 6627
Fax: 519/746-5243
Email: mschulze@uwaterloo.ca