

## Computational Linguistics in Computer-Aided Language Learning: Using NLP for Error Detection and Error Diagnosis

**Course description** Grammar instruction is a necessary component of the foreign language curriculum (cf., e.g., Ellis 1994, Norris and Ortega 2000, Schulz 2002). Nevertheless, the emphasis on communicative language teaching means that only little time is available in the classroom for explicit grammar instruction and feedback on grammar errors. It is therefore attractive to explore whether intelligent tutoring systems can be used to fill this gap by providing individualized feedback to errors made by learners as they progress through sets of on-line exercises.

The seminar will focus on the use of computational linguistics to detect errors in the written output of language learners and generate appropriate feedback. We will look at approaches using deep and shallow processing techniques, with an emphasis on the approaches performing deep analysis based on current constraint-based grammar architectures (HPSG, e.g., Heift 1998; LFG, e.g., Reuer 2003).

- Ellis, N. (1994). Implicit and Explicit Language Learning - an Overview. In *Implicit and Explicit Learning of Languages*. San Diego, CA: Academic Press.
- Heift, G. D. (1998). *Designed Intelligence: A Language Teacher Model*. PhD thesis. Simon Fraser University.
- Nerbonne, J. (2003). *Natural Language Processing in Computer-Aided Language Learning*. In *The Oxford Handbook of Computational Linguistics*, edited by Ruslan Mitkov. Oxford University Press.
- Norris, J. and L. Ortega (2000). *Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-Analysis*. *Language Learning*.
- Reuer, V. (2003). *Error Recognition and Feedback with Lexical Functional Grammar*; *CALICO Journal* 20(3) 2003, 497-512.
- Schulze, R. A. (2002). *Hilft es die Regel zu wissen um sie anzuwenden? Das Verhältnis von metalinguistischem Bewusstsein und grammatischer Kompetenz in DaF*. *Die Unterrichtspraxis-Teaching German*, 35, 1 (2002), pp. 15-24.

## Organizational things

- When and Where: Tuesdays and Thursdays, 9<sup>30</sup>–11<sup>18</sup>, 202 Converse Hall  
Revised schedule proposal:
  - Tuesdays 10<sup>18</sup>–11<sup>18</sup> (202 Converse)
  - Thursdays 10<sup>18</sup>–11<sup>18</sup> (202 Converse) and 5<sup>30</sup>–7<sup>00</sup> (201 Oxley)
- Course email list: 795k@ling.osu.edu
- Course web page: <http://ling.osu.edu/~dm/04/winter/795K/>
- Office hours: Tuesdays, 11<sup>30</sup>–12<sup>30</sup>, 201a Oxley Hall

**General perspective** It is important to emphasize that since this course is a seminar, i.e., each participant is expected to take an active role as a researcher. More concretely, each participant is expected to

- a) explore and present one of the topics:
  - research the topic, starting from (but not ending with) provided references
  - discuss topic with class using handouts, the beamer, or overheads
- b) actively participate in the class discussion, and collaborate as a group. For this it is crucial to take the reading for each session seriously, independent of whether you lead the discussion or not.
- d) explore a handful of exercises on the practical side of the topic
  - Checking out the pages/systems others have produced.
  - Developing increasingly complex web-sites with forms to present material and respond to input.
- e) researches and specify (but not implement) a system providing foreign language learner feedback, which concretely means:
  - explore language learning exercises for language and topic of your choice
  - define a sequence of exercises
  - list envisaged responses and potential errors made by users
  - specify corrections that the system should give
  - discuss technology that could be used to provide this feedback

## The topic

- Background on Computer-Aided Language Learning and Computational Linguistics
  - Nerbonne (2003) and Veit Reuer (2003): “Intelligent Computer-Assisted Language Learning”. Milca project course reader.
  - background on existing systems
- Focus of the seminar: Providing automated feedback to learner answers to exercises.
  - pedagogical aspect:
    - \* explicit instruction and whether it’s useful
    - \* exercises
      - What types of exercises are there?
      - What is which type good for?
      - What errors are made?
      - What knowledge is necessary to provided feedback?
  - CL aspect: Ways to analyze student input and provide error feedback.
    - \* Starting point: Special Issue “Error Diagnosis and Error Correction in CALL” edited by Trude Heift and Mathias Schulze. CALICO Journal. 20 (3). 2003.
      - Mathias Schulze: Errors and Feedback: Some Theoretical Insights
      - Ron Cowan, Hyun Eun Choi, Doe Hyung Kim: Four Questions for Error Diagnosis and Correction in CALL
      - Sylviane Granger: Error-tagged Learner Corpora and CALL: A Promising Synergy
      - Sébastien L’haire and Anne Vandeventer Faltin: Error Diagnosis in the FreeText Project
      - Veit Reuer: Error Recognition and Feedback with Lexical Functional Grammar
      - Rodolfo Delmonte: Linguistic Knowledge and Reasoning for Error Diagnosis and Feedback Generation
      - Trude Heift: Multiple Learner Errors and Feedback: A Challenge for ICALL Systems
      - Cornelia Tschichold: Lexically Driven Error Detection and Correction
      - Liang Chen and Naoyuki Tokuda: A New Template-Template-enhanced ICALL System for a Second Language Composition Course
  - technology, practical skills aspect:
    - \* technical components and realization of web-based presentation of exercises, collection of input, producing and presenting feedback, using python