Chapter 25
Evolution by Evaluation

Cerstin Mahlow
University of Zurich, Switzerland

Michael Hess
University of Zurich, Switzerland

Sven Grund
University of Zurich, Switzerland

ABSTRACT
This paper describes the process of formally evaluating an E-Learning system that has been in use for several years. Professional usability evaluation offers deeper insight into user behaviour and needs than accidental feedback collection or introspection by system developers. A first analysis of the evaluation samples shows satisfaction of users with the general design of the system but also dissatisfaction with certain aspects of navigation that would otherwise have escaped our attention. State of the art formal evaluation turned out to be instrumental in making an existing system considerably more user-friendly.

INTRODUCTION
Students of Computational Linguistics tend to have very different backgrounds: They come from language studies, computer sciences, philosophy, or from various other fields. This heterogeneity is a real challenge for teachers. Part of the audience needs to be given more background about computer sciences while others need to be updated on fundamentals of linguistics, and others yet need to know more about principles of software engineering. Adaptive and interactive E-Learning systems are a good vehicle to help students fill in gaps in their knowledge ahead of lectures and at their own pace. This is why, several years ago, the Institute of Computational Linguistics of the University of Zurich decided to complement all lectures, tutorials and seminars with E-Learning components. Among them is CLab (http://www.cl.uzh.ch/CLab), a web-based virtual laboratory of computational linguistics, used to give students the opportunity to familiarize themselves, in self-study sessions, with the basics of the various fields involved.
CLab has been under constant development for several years, while it was in day-to-day use. In 2005 and 2006 the Institute of Computational Linguistics was a partner in a project, TransTech – Language Technology for Translators (http://www.virtualcampus.ch/display.php?lang=1&pid=239), lead by the École de traduction et d’interprétation at Geneva University. One of the goals of TransTech was to design and then formally and systematically evaluate a number of new CLab modules. We used this opportunity for a general evaluation of the entire CLab as developed so far. In this article we describe prerequisites and preparations for this evaluation, its implementation, and a first data analysis. We will then show the steps to improve CLab, and general lessons learned about the evolution of E-Learning systems.

FUNDAMENTALS OF CLab

CLab offers several units on foundations of computational linguistics and a few on techniques of computer sciences. Each unit consists of an explanatory text, covering one topic in detail. The text is available in two formats: As one large coherent PDF document, and as snippets of HTML text, tightly coupled with the various other components of the unit. Both versions, whose content is identical, contain links to additional texts, to other knowledge resources and, most importantly, to interactive elements. These interactive elements are quizzes to test whether learning goals are achieved, Sentence Completion Tests (SCT) (Mahlow & Hess, 2004) for training and assessment, and Interactive Learning Applications (ILAP) for interactive experimentation with programs (Carstensen & Hess, 2003). In each unit, students are stimulated to interactively explore specific problems, methods, and techniques.

CLab has a clear design. Figure 1 shows a screenshot for the ILAP Chunking before the evaluation took place. In the header you see all navigation information: A link to the CLab start page on the left, the title of the current application (SCT, QUIZ, ILAP) in the middle, help (context help, link to our glossary of computational linguistics) and links to material outside CLab (home pages of the Institute of Computational Linguistics and of the University of Zurich) on the right. The space below is used as working space, divided in columns. The leftmost column is used for instructions and actions user can perform by clicking on buttons. The columns on the right are used for system output (e.g. hits, percentages, reformatted or produced text).

EVALUATING CLab

When we decided, in April 2006, to perform a professional evaluation we knew, on the basis of user feedback, that the system was technically stable. But we were uncertain about various navigation and usability aspects, which is related to Kirkpatrick’s evaluation step 1 “learners satisfaction” (1987). The questions that had to be answered are:

- Do students prefer introductory texts in form of one large, coherent PDF document, or as small HTML pages, as recommended by many web usability studies (e.g. Nielsen, J., 1997)?
- How do students work with the various CLab elements? Do they first read the text as a whole and then work through the interactive elements, or do they start with interactive elements and then consult the text?
- How do students try to achieve specific results with an interactive element in order to answer questions that arose in face-to-face lectures?

We decided to use student feedback and observation of student behaviour to guide our attempts to improve CLab (see Mumford, 1984). In the following we describe in detail the three steps of the evaluation: