Separating the chaff from the wheat:

Creating evaluation standards for web-based language training resources

Lars Borin* and Sara Gustavsson**

** Department of Linguistics, Uppsala University, SWEDEN*
*** Division of IT Services, Mid Sweden University, SWEDEN*

lars.borin@ling.uu.se
sara.gustavsson@forv.mh.se

Abstract*:*

We present the LingoNet project for creating a 'web-based language lab', a website where resources for web-based language training will be collected and made available for use in foreign language education at the university level. One of the most pressing needs in this connection is to develop guidelines, procedures, and tools for the (summative and formative) evaluation of such resources. An important goal of the LingoNet project is consequently to produce such evaluation guidelines. Metadata markup will be used to ensure that information about the resources, including the results of their evaluation—both summative and formative—will be persistent and thus will be fully available to future users of the resources.

Keywords: Language Learning, Learning Resources Evaluation, New Educational Technology, System Design, Web Based Learning

1. Introduction

LingoNet is a one-year R&D project funded by the Swedish Agency for Distance Education (DISTUM; see <http://www.distum.se>). The project is a cooperation between the Divison of IT Services and the Department of Humanities, Mid Sweden University, and the Department of Linguistics, Uppsala University. At the time of the writing of this paper (October 2000), the project has been running for a little over a month.

The aim of the project is to build a 'language lab on the internet', i.e. a web site with a collection of language training resources to be used in higher education, both locally and in distance education. In this connection it is significant that the main project partner, Mid Sweden University, is a 'distributed university', a not uncommon model for newer universities in sparsely populated Sweden. Mid Sweden University has four campuses, in four cities (Härnösand, Örnsköldsvik, Östersund, and Sundsvall), and the distance, as the crow flies, between the two campuses which are farthest apart—Östersund and Örnsköldsvik—is about 200 kilometers. Some departments are located in one of the campuses only, while others are distributed just like the university itself. Among the latter is the Department of Humanities, one of the LingoNet project partners, with parts in both Härnösand and Östersund. Depending on—among other things—the particular program they are enrolled in, students at the university take their courses at one or at more than one of the four campuses. The latter alternative may involve moving to the other campus city, or simply going there regularly (train and bus time tables are among the more prominently displayed information items on the university's website: <http://www.mh.se/>). In any case, the 'local' education at Mid Sweden University contains a fair amount of what elsewhere would be called "distance education". Uppsala University is a larger and more traditional university (Sweden's oldest university, as a matter of fact). Although it, too, offers a fair amount of distance education, the emphasis here is on how IT support can be integrated into local language education. Since Uppsala offers courses and programs in a much wider range of languages than Mid Sweden University (see below), it becomes natural to keep in mind the generality and extensibility of the methods, tools, and resources that we develop in the project.

Even though the point of departure for the LingoNet project is the traditional language lab, we actually envision a more general language training resource than this, i.e. a 'computer language lab', rather than a 'computerized version of the tape-based language lab', as the idea is not only to transfer older techniques into this new technology, but also to exploit the additional possibilities offered by the new technology itself.

Anybody who has looked for language learning sites on the web knows that there are quite a few such sites, including sites with large numbers of links to various kinds of resources of the kind which could make up a language lab. What, then, is the point of creating yet another such site? Could we not simply point our students to some such sites, or even ask them to find appropriate sites for themselves?

When we ask of somebody that they "look for information on the web", or "find resources on the internet", we are actually expecting of them that they have mastered the highest and most difficult competence in Bloom's taxonomy of educational goals [BLO 56], that of being able to evaluate information. This means, however, that students are actually ill-equipped to use raw web content for their education, in any field, because they do not in fact yet possess the necessary skills to evaluate this content, these being skills which they hopefully will acquire in due

course, as a result of the education they are undergoing, but which cannot be presupposed.

Our academic teachers, on the other hand, who are perfectly capable of evaluating this raw web content, are instead extremely pressed for time, and surprisingly often still (at least in the humanities) less than comfortable using computers, let alone learning how to find[4], how to use, and possibly how to adapt new computer applications to their needs.

In the LingoNet project, we have seen a possible way out of this catch-22 situation through the use of *quality control* and *metadata*. It is a well-known fact that the information to be found on the web on any topic is, not only abundant in almost all cases, but also—to put it mildly—of extremely varying quality. At the same time, web search engines are still fairly primitive, so that finding educational resources, appropriate as to their content and level—regardless of their quality—in itself takes some work [HOW 99: 24f]. It is only after they have been found that the real work begins, however, when the chaff—resources which are of low quality or of the wrong kind—is to be separated from the wheat—the resources which we can use for our educational purpose, i.e. educational web resources which are quality controlled and classified as to their content and level.

As we have already indicated, one of the main aims of the LingoNet project is to collect and if needed create such quality controlled and classified resources for language training at the university level. In this case it is natural that the quality control takes the form of *evaluation*. Evaluation traditionally comes in two forms, *summative* and *formative*, and we see natural uses for both kinds in the LingoNet project. More will be said about this in the following sections.

The difference between summative and formative evaluation lies mainly in how the results of the evaluation are meant to influence the development of the application undergoing evaluation. In the case of summative evaluation, no particular such influence is foreseen. In summative evaluation, an application's appropriateness for a particular goal is assessed, and the outcome of the evaluation is simply a summary of its strong and weak points, and possibly a recommendation to use or not to use the application for this goal. Just like summative evaluation, formative evaluation also assesses the application's appropiateness for a goal, but in this case the evaluation is seen as part of the application development cycle, and the results are used for improving the application. In the software industry, usability lab testing and so-called beta testing are frequently encountered formative evaluation methods.

---

[4] Whereas academic language teachers can safely be assumed to be perfectly capable of evaluating digital language learning resources, wielding web search engines to locate the same resources is an ability which *cannot* be taken for granted in this population.

2. Classification and evaluation of language training resources

2.1 Collecting resources

The first phase of the classification and evaluation of web-based language training resources will be their inventory. This is work on which we have already started in the project. The inventory will never be complete, of course, because the WWW is so vast that we can have no real hope of finding all the relevant resources present there at a particular point in time. More importantly, however, the web is also growing quickly, so that the inventory becomes something of a moving target. However, anybody who has done background research in preparation for writing a scientific paper knows that this is nothing new to the WWW, only the scale of it is, and that in both cases the quality of the (re)sources is actually more important than their quantity.

The reason that we say that the resource inventory comprises the first phase of our work, rather than forming a prerequisite for it, is that we recognise that it should not be conducted as a kind of blind trawl for resources, as this in our view is not the proper way to promote quality. Rather, the inventory should be guided by an informed view of what a web-based language lab *should* look like, i.e. which kinds of resources should make up the language lab, and how these resources should work together and fit into the university language educational setting. This, however, is a question which is difficult to answer in the abstract, and which actually *should not* be answered in the abstract. Consequently, the point of departure for our search for resources will be a 'wish list', a specification which language teaching faculty at the Department of Humanities, Mid Sweden University is compiling. This means that their very concrete needs will be the main factor determining the final structure and functionality of the web-based language lab. Even more concretely, this specification will be limited to the three languages English, French and German (as foreign languages), which are the ones currently taught in the Department of Humanities. Certainly at least part of it will generalize readily to the teaching of other foreign languages as well, which is most desirable (see above), but equally certainly there will be another part which is more narrowly applicable to the three mentioned languages and similar languages only (e.g., to Spanish or Dutch, but not to Quechua or Frisian). The work on the specification started in September 2000. While still at a preliminary stage, it has already resulted in a coarse provisional structure for the LingoNet language lab, with three main link categories (reference materials, skills training resources, and others—the proposal actually calls this last category "junk links") and some subcategories.

The compilation of a 'wish list' of this kind does not do away with the need for a more general search for resources, however. It only means that the search can be more focussed. Firstly, the needs expressed by the educators can be vague and result-oriented (as opposed to detailed and means-oriented), and secondly, the

wishes that they will be able to come up with will perhaps more often than not be influenced by what they know or believe will be possible to bring about. The educators should be allowed to be vague and result-oriented in their wishes, and it should be our task to come up with suggestions for applications which fit these wishes. For the second point, in order to get the educators not to set their sights to low, we should have carried out a preliminary mapping of the space of language-learning applications on the web prior to the compilation of the wish-list, and we should also be able to give them some idea of the state of the art of the field. We are at present in the process of collecting and compiling material on both these topics. In order to streamline this process, a website has been set up at Mid Sweden University, where project members can add links to language learning resources that they have found on the web, together with (free text) comments about the resources.

The resources will be evaluated for possible inclusion in LingoNet during the second phase of the project (see the next section). The same website which is at present used for collecting resources, as explained in the previous paragraph, will be used for classifying and sorting them. The teachers will now be able to go through the resources and add structured annotations (as opposed to the free text comments added earlier) about their purpose and suitability for that purpose. If there is a need for a kind of resource which is not generally available, then, depending on the circumstances, we might decide to develop this resource in the project (see section 2.4, below).

2.2 Summative evaluation of resources

The second phase of the project will yield two kinds of result. Firstly, there will be a set of quality-controlled web-based language training resources, arrived at through a systematic summative evaluation of the resources located in the first phase. Secondly, there will be a set of guidelines for carrying out this kind of evaluation in the future. The latter result, in our view, will be one of the most important outcomes of the LingoNet project, and the one which inspired this paper. Its importance is mainly due to the circumstance that, if the LingoNet concept turns out to be a good way of getting faculty and students to use high-quality, effective web-based language training resources, the guidelines will be the key—*mutatis mutandis*, of course, but hopefully with most of the ground covered already—to building other similar sites, as well as to keeping the LingoNet site continuously updated with new applications. If it is to have this kind of lasting effect, however, the set of guidelines should focus on less ephemeral aspects of the resources, i.e. the emphasis should be on such factors which can be assumed to stay relatively constant even with rapid changes in the underlying technology. The evaluation guidelines should thus—and this is nothing new or surprising; we are merely stating it for the record, as it were—emphasize content and effects, rather than technology.

Although we will be building on existing guidelines for the development and evaluation of applications for computer-assisted language learning [CAL 00; CAM 90; CHA 97; CHA 98; COL 96; COU 96; HEW 98; HUB 96; LAU 96; MB 98;

WÅL 00], computer-assisted learning [ARM 99; COL et al 97; LAT 99], and for the design of user interfaces and related matters [SHN 98; CAR et al 99], we will certainly need both to combine and add more detail to various parts of these guidelines. This is because existing guidelines have been developed for many purposes and many kinds of learning (resources). For instance, it seems that there have been relatively few attempts to evaluate language learning resources for higher education, as opposed to the primary and secondary school levels. Also, there are fairly detailed 'checklists' for evaluating multimedia programs for communicative language learning activities, but much less material on evaluating e.g. corpus processing tools used in computer-assisted language learning [FLO 96; KNO 90].

The general procedure at this stage will be one where educators from the Department of Humanities, Mid Sweden University, evaluate each resource according to the (simultaneously evolving) evaluation guidelines. As a first approximation of the final categorization, each language learning resource will be classified according to language, level of difficulty, resource type and how relevant it is deemed to be for the language lab. At least the following general aspects will be covered in the evaluation process, and thus also in the guidelines, but the list is still evolving, and we will need to flesh out the general aspects in order to arrive at the level of detail necessary for our purposes. In connection with this, we need to resolve the issue whether we are developing a kind of *checklist* or a *set of rubrics* (although at a finer level of detail than that shown below) [LAT 99: Introduction].

- *content* — e.g. type and level of difficulty of exercises
- *pedagogy* (subject-specific and general) — underlying pedagogical model, error handling etc.
- *user interface* — how well the interface expresses the pedagogical idea, navigation, ease of use, feedback provided etc.
- *documentation* (in a wide sense) — are there supplementary materials or instructions for students and teachers, or documentation on how to adapt the resource to particular needs?
- (predicted) *acceptance* by faculty and students — is the program likely to be used and appreciated by the intended users?
- *reusability* (and generality) — can the resource easily be extended with more material, new exercises be added, etc.?
- (predicted) *efficiency and effectiveness*
- *technology*
- *predicted learning outcomes* — what skills will the student acquire? Here Bloom's Taxonomy of Educational Objectives [BLO 56] can be used to identify precise and measurable definitions of learning outcomes, i.e. is recall, comprehension, application, analysis, synthesis or evaluation (judgment) involved?

As always when classificatory schemes meet real life, there is overlap and dependencies between these aspects. For example, the "(predicted) acceptance by faculty" will to some extent be dependent upon how well the user interface answers to the needs and abilities of our academic teachers, how good the documentation is, etc.

On the basis of the evalutation, the decision will be made to incorporate or not incorporate each resource in the LingoNet website. The results of the evaluation will also become part of the markup of the resource (see the next section).

2.3 Classification and marking-up of resources

The third phase of the process will consist in the classification and marking-up of resources, with both content descriptions and evaluative statements. Thus we ensure that future users will have access to this information. In order to make the classification maximally accessible, we propose to use the emerging Learning Objects Metadata (LOM) standard developed jointly by the IEEE Learning Technology Standards Committee (<http://ltsc.ieee.org/>), the Instructional Management System project (<http://www.imsproject.org/>), the European ARIADNE project (<http://ariadne.unil.ch/>) [IEE 00; AW1 00; AW2 00], and other organizations, probably in its XML form [AW3 00].

The Learning Objects Metadata information model [AW1 00] specifies metadata elements in nine areas:

(1) general: "Context-independent features of the resource." [AW1 00: 6]
(2) lifecycle: "Features related to the life cycle of the resource." [AW1 00: 9]
(3) metametadata: "Features of the description rather than the resource." [AW1 00: 10]
(4) technical: "Technical features of the resource." [AW1 00: 12]
(5) educational: "Educational or pedagogic features of the resource." [AW1 00: 14]
(6) rights: "Conditions of use of the resource." [AW1 00: 17]
(7) relation: "Features of the resource in relationship to other resources." [AW1 00: 17]
(8) annotation: "Comments on the educational use of the resource." [AW1 00: 18]
(9) classification: "Description of a characteristic of the resource by entries in classifications." [AW1 00: 18]

Each area contains a number of core and extension elements. Broadly speaking, *classification* is better catered for than *evaluation* in this metadata scheme, in that almost all nine areas can be said to have at least some classificatory content.

At any rate, both the classification and educational metadata areas have as their primary function that of classification of the resource in an educational setting. Evaluative statements as well as the results of evaluation can presumably go only under the heading of annotation ("Comments on the educational use of the resource."), which contains three kinds of (extension) elements: a person, a date, and a description (i.e., the annotation itself). Hence, there seems to be an asymmetry in the way classification and evaluation are handled in the LOM. It provides a fine-grained predetermined classificatory scheme, but for evaluation there is only a single kind of free text element. We have yet to decide whether to introduce further structure into this element. The advantage of such a move would be that structured information is generally more useful than unstructured data. The drawback, on the other hand, would be loss of generality; basically, we would have a metadata element which only our search engine would be able to parse. Perhaps a compromise can be reached where the annotation.description element—i.e., the description element of the annotation area—is a free text string in a so-called controlled language. You would still need a parser for this controlled language in order to be able to reap the full benefits of the evaluation data, but it would also be understandable as free text. We are still considering this issue, however.

For resources that we develop in the project (see section 2.4.1), we aim to use LOM throughout (possibly in addition to other kinds of more resource-specific metadata), but in the case of resources which already exist, the metadata will reside on the LingoNet website, together with a hyperlink to the resource (found in the element technical.location). Hence, in the latter case, LingoNet provides a classification for an existing resource, but the classification will not be available in the resource itself.

LingoNet will thus make its language training resources retrievable on the basis of LOM metadata. We have not yet decided whether there should also be a general LOM metadata based search engine for resources not formally incorporated in LingoNet. On the one hand, this would go against the spirit of the project (at least in the narrow sense). On the other hand, one could perhaps assume that resources which have been specifically marked-up with LOM metadata have also gone through some kind of evaluation process, so that this would not be the same case as that of finding a putative resource by a general keyword search. For the reasons given above, the results of such an evalution may not be accessible, but the classification information should be.

In short, the function of LOM metadata markup is primarily to make it easier to find an educational resource of a particular kind, or to fit a newly acquired educational resource into an existing digital resource structure (i.e., an important function of metadata markup is to enhance resource portability), but we will also make use of it to make the result of the summative evaluation persistent and generally available.

2.4 Formative evaluation of resources

The formative evaluation which we foresee in the project will actually be of two kinds: 'real' formative evaluation—the topic of the next section—and 'continuous' formative evaluation, which will be discussed in section 2.4.2.

2.4.1 'Real' formative evaluation

The first kind of formative evaluation will be used for the resources which possibly will be developed in the project to fill an expressed educational need for which no resources are to be found on the web (see section 2.1). The decision whether to develop new resources will be based on how deeply felt the need for this particular resource is, but also on the expertise and funding needed for the development effort. If it is decided to develop a particular resource, this development will be carried out by the Division of IT Services at Mid Sweden University in cooperation with the Department of Linguistics at Uppsala University. The formative evaluation work will be done partly in the Department of Humanities and partly in the Usability Laboratory of the Division of IT Services at Mid Sweden University.

2.4.2 'Continuous' formative evaluation

The other kind of formative evaluation will be an integral part of the LingoNet website itself. We plan to build into it facilities for continuous feedback from the users of the LingoNet resources, mainly students and faculty in higher education, but others will be able to contribute their views and suggestions as well. In other words, we wish to build in evolvability in the website. Not only direct feedback, but also usage patterns could be recorded and utilized for formative evaluation purposes, e.g. by the use of data mining techniques. This is a matter which we are planning to look into, but which strictly speaking falls outside the scope of the present project, and thus must be given a lower priority at present. It will clearly be a very natural goal for a possible follow-up project to LingoNet, however, and we hope to be able to return to this matter in the not-too-distant future.

Just as for the summative evaluation, there will be a need to develop guidelines and easy-to-use procedures for continuous formative evaluation as the LingoNet website itself takes more firm shape. This again is of somewhat lower priority than the guidelines for summative evaluation, but something that would make up a natural and important research topic in a follow-up project. Needless to say, there is considerable overlap between the formative and summative evaluation modes, so that much of the effort which will have gone into the development of guidelines, procedures and tools for summative evaluation, will indubitably turn out to be of use also in the context of formative evaluation.

3. Summing up

One of the most important goals of the LingoNet project is in our view the creation of a set of guidelines for summative evaluation—and to some extent also for formative evaluation—of web-based language training resources suitable for use in foreign language education at the university level. These guidelines will be based on concrete educational needs and will be the result of an continuous interaction between language educators in the Department of Humanities, Mid Sweden University and the LingoNet website developers in the Division of IT Services, Mid Sweden University and the Department of Linguistics, Uppsala University. The guidelines will emphasize content and effects of the resources, rather than their underlying technology. They will further facilitate the construction of similar websites in the future, as well as the continuous updating of the LingoNet site.

An equally important goal of the project is that of using the emerging educational technology LOM metadata standard for annotating the LingoNet website resources, both with regard to their content and purpose, and with regard to the result of (1) summative evaluation made by language educators at Mid Sweden University, and (2) formative evaluation made by users—both students and teachers—of the LingoNet website.

One aspect which we have not touched upon in this paper, but which is of paramount importance for the ultimate fate of the LingoNet website, is that of how the resources aviailable there are integrated into existing course structures, alternatively—which is the more reasonable scenario—how existing course structures are modified to accomodate web-based language training resources. Presumably, the educators evaluating the resources will need to look into this issue, but the proof of the pudding will ultimately be in whether it is eaten or not, i.e. whether others[5]—teachers and students—will find the LingoNet setup useful in university-level language learning. Since our project budget does not provide for the acquisition of a chrystal ball, only time will tell whether this will turn out to be the case.

References

[ARM 99] Armstrong, Sara 1999. A Framework for Evaluating Integrated Learning Systems/ Multimedia Instructional Systems. California Instructional Technology Clearinghouse. Retrieved from the WWW in August 2000 via <http://clearinghouse.k12.ca.us>.

[5] Others than the project members, that is. The project members are in the project at least partly because they are technogically more literate than some of their colleagues, and excited about the possibilities of using online educational resources in their teaching.

[AW1 00] Anderson, Thor and Tom Wason 2000. IMS Learning Resource Meta-data Information Model. Final Specification. Version 1.1. Retrieved from the WWW in July 2000: <http://www.imsproject.org/metadata/mdinfov1p1.html>.

[AW2 00] Anderson, Thor and Tom Wason 2000. IMS Meta-data Best Practice and Implementation Guide. Final Specification. Version 1.1. Retrieved from the WWW in July 2000: <http://www.imsproject.org/metadata/mdbestv1p1.html>.

[AW3 00] Anderson, Thor and Tom Wason 2000. IMS Learning Resource XML Binding Specification. Final Specification. Version 1.1. Retrieved from the WWW in July 2000: <http://www.imsproject.org/ metadata/mdbindv1p1.html>.

[BLO 56] Bloom, B.S. (ed.) 1956. *Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook I, Cognitive Domain*. New York: Longman.

[CAL 00] CALICO 2000. Software Evaluation Outline: Courseware. Retrieved from the WWW in August 2000: <http://astro.temple.edu/~jburston/CALICO/softrev.htm>.

[CAM 90] Cameron, Keith (ed.) 1990. *Evaluation of CALL Programs* [= *Computer Assisted Language Learning*, Vol. 1]. Oxford: Intellect.

[CAR et al 99] Card, Stuart K., Jock D. Mackinlay and Ben Shneiderman (eds.) 1999. *Readings in Information Visualization. Using Vision to Think*. San Francisco: Morgan Kaufmann.

[CHA 97] Chapelle, Carol 1997. CALL in the year 2000: Still in search of research paradigms? *Language Learning and Technology* 1(1). <http://polyglot.cal.msu.edu/ llt/>, p 19–43.

[CHA 98] Chapelle, Carol 1998. Multimedia CALL: Lessons to be learned from research on instructed SLA? *Language Learning and Technology* 2(1). <http://polyglot.cal. msu.edu/llt/>, p 22–34.

[COL et al 97] Coley, Richard J., John Cradler and Penelope K. Engel 1997. Computers and Classrooms: The Status of Technology in U.S. Schools. ETS Policy Information Center Report. Princeton: ETS. Retrieved from the WWW in July 2000: <ftp://etsis1. ets.org/pub/res/compclss.pdf >.

[COL 96] Colpaert, Jozef 1996. Learning from the past, building new working hypotheses: The Didascalia criteria framework for more added value in CALL. *Computer Assisted Language Learning* 9(4), p 309–318.

[COU 96] Council for the Renewal of Undergraduate Education 1996. The Reference Group for Promoting the Use of Multimedia in Language Education (LingCAL): Utvärderingsmall ('Evaluation Outline'; in Swedish). Retrieved from the WWW in December 1999: <http://www.hgur.se/lingcal/Kriterier/JFkriterier.html>.

[FLO 96] Flowerdew, John 1996. Concordancing in language learning. *The Power of CALL*, edited by Martha C. Pennington. Houston: Athelstan, p 97–113.

[HEW 98] Hewer, Sue 1998. Optimising the Use of TELL Products: an Evaluative Investigation into TELL Products in Use. The TELL Consortium, University of Hull. Retrieved from the WWW in November 1999: <http:// www.hull.ac.uk/cti/sumeval. doc>

[HOW 99] Howard Chen, Hao-Jan 1999. Creating a virtual language lab: an EFL experience at National Taiwan Ocean University. *ReCALL* 11(2), p 20–30.

[HUB 96] Hubbard, Philip L. 1996. Elements of CALL methodology: Development, evaluation, and implementation. *The Power of CALL*, edited by Martha C. Pennington. Houston: Athelstan, p 15–32.

[IEE 00] IEEE 2000. IEEE P1484.12/D4.1. 12 March 2000. Draft Standard for Learning Object Metadata. Sponsored by the Learning Technology Standardization Committee of the IEEE. Retrieved from the WWW in September 1999 via <http://ltsc.ieee.org/ >.

[KNO 90] Knowles, Gerald 1990. The use of spoken and written corpora in the teaching of language and linguistics. *Literary and Linguistic Computing* 5(1), p 45–48.

[LAT 99] Lathrop, Ann (ed.) 1999. Guidelines for the Evaluation of Instructional Technology Resources for California Schools. California Instructional Technology Clearinghouse. Retrieved from the WWW in June 2000 via <http://clearinghouse. k12.ca.us>.

[LAU 96] Laurillard, Diana 1996. The TELL Consortium – Formative Evaluation Report. The TELL Consortium, University of Hull. Retrieved from the WWW in August 1998: <http://www.hull.ac.uk/cti/formeval.doc>.

[MB 98] Murray, Liam and Ann Barnes 1998. Beyond the "wow" factor—evaluating multimedia language learning software from a pedagogical viewpoint. *System* 26, p 249–259.

[SHN 98] Shneiderman, Ben 1998. *Designing the User Interface. Strategies for Effective Human–Computer Interaction*. 3rd ed. Reading, Massachusetts: Addison–Wesley.

[WÅL 00] Wålstedt, Elin 2000. English On-line – Utvärdering av ett internetbaserat CALL-program ['English On-line – Evaluating an internet-based CALL program'; in Swedish]. Master's thesis in Language Engineering at the Department of Linguistics, Uppsala University.