Requirements on Infrastructure for CALL (Computer Assisted Language Learning)

Robert Bannert[1], Antonis Botinis, Barbara Gawronska, Gerd Hollenstein[2]

Abstract

The present contribution is an outline of the development of CALL applications at the University of Skövde, Sweden. Current pedagogical approaches and teaching methodologies as well as linguistics and language technology are brought together in a unified context. Our CALL system targets to the acquisition of fundamental language skills, in the first place reading, writing, speaking, listening and understanding. The materials consist of written texts red aloud, guided and spontaneous dialogues, and other exercises. Main components of the language, such as vocabulary, pronunciation and grammar are organized in training pathways, in an ascending order, in accordance with the language acquisition requirements. Diagnostic tools as well as repeated evaluation testing are also included in the system.

I. INTRODUCTION

The state of the art of Computer Assisted Language Learning (CALL) applications depicts a promising situation in which pedagogical principles and technological potential are brought together ([1] - [5]). Technology integrates rapidly in major spheres of human interaction and its influence on language education is widely acknowledged. Despite a worldwide increase in CALL applications in recent years, there is still an urgent need for further development and more sophisticated systems to meet a wide range of user needs and education requirements.

The present paper concentrates on methodological principles adapted in a project recently started at the University of Skövde, Sweden – UNICALL (University Computer Assisted Language Learning). Current pedagogical methods in second language education are related to the state of the art of computational linguistics, language technology, and speech technology.

II. CALL AS AN INTERDISCIPLINARY RESEARCH AREA

CALL applications involve a variety of disciplines, from psychology and linguistics to speech and language technology. This is a tradition in which typical CALL models are based. For example, Ahmad [6] summarizes the relations between three main components of a CALL system (the learner, the language and the computer) and the related scientific areas shown in figure 1:
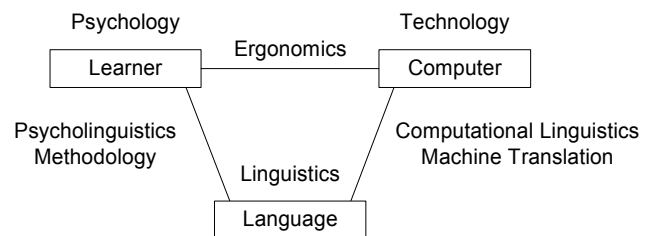

Fig. 1. A typical CALL model and related scientific areas ([6], p.45).

Such models do not however take much advantage of recent developments in technology and language education methodologies. Firstly - because of the general access to Internet and World Wide Web - it would be more appropriate to replace the single computer in Figure 1 by a computer network; secondly, language should not be isolated from cultural knowledge. The last claim implies that such research areas as cultural studies, intercultural communication, sociolinguistics and psycholinguistics should be taken into account. In the framework of the UNICALL project we have developed a powerful interdisciplinary model in accordance with figure 2. Critical aspects of language learning such as the target language (L2) in relation to the native language (L1) as well as the communicative situation and the actual context of the language in question in relation to the actual learner are basic components our model. In addition, theoretical aspects of phonetics, and pronunciation practice and exercises (utilizing speech and language technology tools and methods) are major components of the model.

[1] Umeå University, Sweden, SE-901 87 Umeå, email Robert.Bannert@ling.umu.se
[2] University of Skövde, Sweden, Box 408, 541 28 Skövde, emails: antonis.botinis@isp.his.se, barbara.gawronska@isp.his.se, gerd.hollenstein@isp.his.se
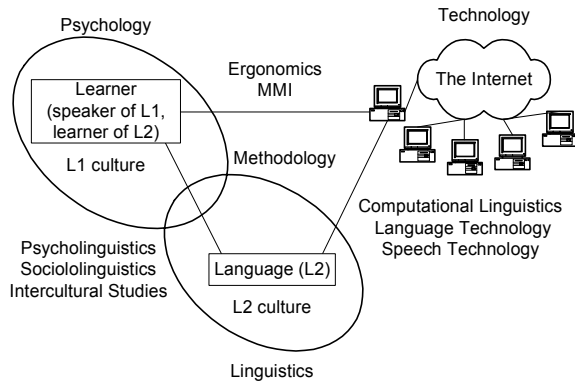
Fig. 2. UNICALL model and related scientific areas

III. FROM MODEL TO IMPLEMENTATION

The first necessary step towards an implementation of the model sketched in fig. 2 is to make decisions with respect to the following aspects:

1) The target user group;
2) User needs: how to diagnose the user's present language skills? What kind of training is needed?
3) Educational methods;
4) Educational resources: linguistic and culture-oriented (general and domain specific lexicons, thesauri, text- and dialogue corpora);
5) Language technology tools (spelling checkers, grammar checkers, taggers, parsers, morphological processors, translation tools, text-to-speech tools);
6) Interface design;
7) Evaluation strategy;

In the following, we concentrate on points 1-5, since an exhaustive discussion of interface design would require a separate paper. Evaluation strategy will be addressed in the last section.

A. Target groups

The target user groups of the system to be developed within the UNICALL framework will primarily be new-enrolled Swedish university students as well as the respective tutors in English, French, German, and Spanish. We also expect that even advanced high school students intending to study these languages at university level will benefit substantially. The system, in the first place, is thought to be appropriate for regular class courses as well as smaller groups and individual use by the Internet. However, the system, due to its open architecture, will in principle be able to develop any language. Language education for specific purposes in various professions is within the broad scope of this system.

B. Analysis of user needs

Since 1991, the Council of Europe has been working on guidelines and tools for measurement of language learners' proficiency. The result is the Common European Framework of Reference for Languages (CEF; [7],[8]). The CEF will be employed in the current project, since it provides "a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe" ([7], p.1), and contains very useful proficiency level definitions (global and detailed), intended to facilitate comparisons between different systems of qualifications. The scheme below shows the proficiency levels adapted by the CEF ([7], p. 23, [9]):
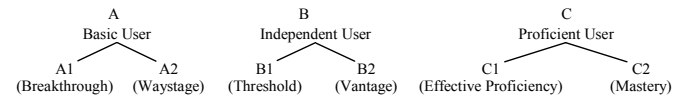
The CEF provides also a tool for classification of learners' skills - so-called "Can Do" descriptors, dealing with language reception, interaction, and production. Below, some of the "Can Do" descriptors for B1 and C1 levels are shown ([7], pp.26-27)

TABLE I
FRAGMENTS OF 'CAN DO' DESCRIPTIONS OF B1 AND C1 PROFICIENCY LEVELS
IN COMMON EUROPEAN FRAMEWORK, PP. 26-27

	B1 (Threshold)	C1 (Effective Proficiency)
Listening	I can understand the main points of clear standard speech on familiar matters (…) I can understand the main point of many radio or TV programmes on current affairs (…) when the delivery is relatively slow and clear.	I can understand extended speech even when it is not clearly structured (…) I can understand television programmes and films without too much effort
Reading	I can understand texts that consist mainly of high frequency everyday or job related language. I can understand the description of events, feelings and wishes in personal letters.	I can understand long and complex factual and literary texts (…) I can understand specialized articles and longer technical instructions, even if they do not relate to my field.
Spoken Interaction	I can deal with most situations likely to arise whilst traveling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work and current events)	I can express myself fluently and spontaneously without much obvious searvhing for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skillfully to those of other speakers.
Spoken Production	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.

An extension of the Common European Framework is the introduction of a European Language Portfolio, intended to

make it possible for learners to document their progress by recording learning experiences. This initiative will be taken into account within the framework of UNICALL.

Since 1997, a project called DIALANG [10] is going on, aimed on making the "Can Do" descriptors available on Internet for quick self-assessment with automatic response. For the time being, the DIALANG service can be used for four languages only, but the versions for all EU-languages are expected to be released within the next few years. DIALANG would provide an excellent basis for an effective diagnosis of the individual user's needs. Even before the automated services for German, French, and Spanish become available, we intend to perform diagnostic tests in accordance to the CEF-policy.

A shortcoming of the CEF diagnostic tools is the lack of possibility of evaluating the user's pronunciation proficiency. A component for diagnosis of the user's skills with respect to certain phonetic features of the target language (especially accentuation patterns and those phonetic distinctions that are crucial for target language understanding – cf. [11]) is therefore to be added in UNICALL.

C. Educational methods

In most general classifications of methods in second language education ([1],[12]) a distinction is made between explicit (deductive) learning, learning based on behaviouristic methods, and implicit learning (language acquisition). An overwhelming part of traditional language teaching material is organized in accordance with the two first mentioned methods. Learners are expected to understand and memorize general grammar rules (for example, word-order rules or declension patterns) before applying them to specific utterances (this is the deductive, or explicit approach). This task is supported by repetitive exercises and reinforcement, i.e. repetition of the correctly learned structures in the next part of the teaching material (the behaviouristic method).

Since about 1970, more and more criticism has been raised against the traditional methods ([13] - [15]). It has been pointed out that the deductive and behaviouristic methods focus too much on accuracy, i.e. on production of syntactically and morphologically correct responses. This accuracy-centered learning often results in "teacher-oriented" language skills, which may turn out to be of no or very little use in authentic communicative situations ([16]). Most educational theoreticians of today (cf.[3],[17]) agree as to the fact that the key notions in successful second language education should be:

1) Active acquisition: the learner shall discover general patterns and rules by studying natural language data;
2) Communicative competence, interactivity: language fluency at least as important, or possibly more important than accuracy; the learning process should be meaning-centered rather than form-centered;
3) Cultural competence (without it, real communicative competence cannot be achieved);
4) Authenticity: the learner should be confronted with authentic texts and dialogues, authentic environments and tasks, communicate with an authentic auditory);
5) Flexibility and dynamics: educational tools and materials should be adaptable for use in different circumstances and capable of further evolution and refinement in response to users' experience);
6) Learner-centered approach (individualized tasks, individualized evaluation; for details, see [3]).

The principles enumerated above constitute the methodological basis of the UNICALL project. It should, however, be kept in mind that certain elements of the deductive or the behaviouristic approach cannot be totally excluded. For example, pronunciation training would be impossible without repetitive exercises. Furthermore, as e.g. Healey ([18]) points out, there actually exist learners "who work well by starting with language rules" as opposed to those who prefer to start with language data, so the access to linguistic generalizations should not be absolutely prohibited.

D. Linguistic and culture-oriented resources

This section is not thought as a survey over existing educational resources, but as a specification of requirements that should be fulfilled in order to meet the target groups' expectations and to implement the methodological approach outlined in section III C. Up to now, our investigations of available resources indicate that existing electronic corpora and dictionaries require a lot of improvement, additions and re-structuring. For example, the large lexical database for English, WordNet ([19],[20],[21]) is a very outstanding linguistic resource due to its sophisticated organization (the lexical items are connected by a number of linguistic relations, like synonymy, antonymy, superordination etc.). Nevertheless, several improvements are needed in order to make it work in an automated application, especially with respect to the problem of word ambiguity ([22],[23]). Furthermore, the database does not contain pronunciation instructions. On the other hand, there exist very good pronunciation dictionaries that lack semantic information and semantic connections between lemmas.

The ambition of the UNICALL project is to exploit the existing resources and to develop re-usable components and a suite of modular building blocks. The result is intended to comprise resources with the following characteristics:

1) Corpora: authentic texts and dialogues, inclusive video-recorded dialogues, grouped into topical groups covering everyday life situations, like travelling, shopping, eating out, etc. and different cultural aspects (environment, health care system, educational system, administration, insurance system, politics).
2) Multimedia dictionaries should encompass the following features:
   a) pronunciation of the words by native speakers

representing different language variants (e.g. Australian, Canadian and American English);
b) grammatical information;
c) presentation of the word usage within real contexts;
d) definition of the word in the same language;
e) domain information (e.g. medical science, chemistry, politics…);
f) translation equivalents;
g) audio-visual material related to the semantic content of the word;
h) intelligent search mechanisms (searching for synonyms, semantically related words, part-of-the-word search);

3) Exercises closely related to the corpus material and aimed at training a wide variety of language phenomena (for a sample exercise, see section E 1).
4) User's database, enabling users to create their own texts/dialogues and their individual dictionaries, and making it possible to integrate videos, sounds and images.

Those components cannot, however, replace an authentic audience. The teachers will therefore be encouraged to complete the system with real interactive tasks (e-mail- or chat-communication with native-speakers and learners of the target-language).

E. Language technology tools and integration of tools and resources – an example

A detailed specification of the requirements to be put on such different language technology tools as grammar checkers, morphological analysers, translation tools, parsers etc. would exceed the scope of the current paper, and make the presentation far too technical and specialized. Instead, we will try to show how different resources and tools may be integrated within a single multi-purpose language exercise.

1) A fragment of a dialogue exercise

The purposes of the exercise are: to promote understanding of spoken German, to extend the vocabulary within the domain of travelling, especially "documents needed during a travel", and, finally, learning certain German syntactic constructions that differ structurally from their Swedish translation equivalents.

1) Start: the user is presented a video sequence, showing a policeman that stops a car driver.

Policeman: Guten Tag. Verkehrskontrolle. Ihren Führerschein und den Zulassungsschein, bitte. ('Good day. Traffic control. Your driving licence and the car registration, please')
Driver: Ich glaube, die habe ich zu Hause liegen lassen ('I'm afraid, I've left those at home.')

If the learner understands the dialogue in its spoken form, the video continues. Otherwise, the user may stop

the video sequence and see the dialogue in written form.

2) Suppose that the learner does not understand the word *Zulassungsschein* ('car registration'). Clicking on the word results the following help alternatives, the help tasks to be performed by resources/tools shown below:

TABLE II
HELP OPTIONS FOR SINGLE WORD UNDERSTANDING

Help option	Result	Resources/tools involved
Morphological analysis and synonyms of word parts	Zulassung = Genehmigung, Erlaubnis Schein = ein offizielles / amtliches Dokument, Zeugnis	Morphological analyser, semantic dictionary with domain information
Synonyms	Fahrzeugsschein	Semantic dictionary
Show picture	Picture of a car registration document	Multimedia dictionary
Definition in German	Ein Dokument mit den Daten von Fahrzeug und Besitzer	Semantic dictionary
Translation	"Registreringsbevis"	Links between the German and the Swedish dictionary

3) Suppose the learner is not familiar with the construction *die habe ich zu Hause liegen lassen*. ('I've left those at home', lit.: 'I have those at home lay leave')

Marking the whole phrase gives the following help alternatives:

TABLE III
HELP OPTIONS FOR SENTENCE UNDERSTANDING

Help option	Result	Resources/tools involved
Syntactic analysis	Subject (who): ich Predicate (has done): habe - liegen lassen object (what): die adverbial (where): zu Hause	Parser (syntactic analyser) Grammatical information from the dictionary
Paraphrase	Die sind zu Hause Die liegen zu Hause (they are at home)	Semantic analyser Sentence generator
Translation	Dem har jag (lämnat) hemma	Machine translation tool

4) The learner has now – using the above exemplified help tools – understood the dialogue. He/she can now proceed to open dialogue exercises within the domain "travel and control" (for example, the next video shows a passenger on a train and a guard,

and the learner can play one of the parts in a new dialogue (*Guten Tag. Fahrkartenkontrolle. Ihre Fahrkarte bitte. / Ich glaube, die habe ich…*). The learners' input is checked by morphological, syntactic and semantic analysers, all of them having access to the dictionary.

Implementation of this kind of exercises is fully realistic given the state of the art of computational linguistics and language technology of today. The most important technical requirement is flexible and domain-related interaction between different tools and resources.

2) Pronunciation exercises

A remaining problem is the difficulty of processing spoken language input. The existing speech recognition tools are yet not good enough to be incorporated in a language exercise of the above outlined kind. Speech recognition tools utilized today (for example, in the speech-to-speech translation system Verbmobil – [24]) work relatively well on standard language, but have problems with processing speech with (even very slight) foreign accent and can therefore not be utilized for language education.

There are, however, possibilities of using speech technology tools for pronunciation training. The learner can, for example, record his/hers utterances (single words, or phrases, or whole sentences) and compare their phonetic and prosodic features with the native pronunciation by means of spectrograms and intonation curves. Figures 3 and 4 show an example – the difference between the correct German pronunciation of the word *Technik* (with stress on the first syllable – fig.3) and a pronunciation error very frequently made by Swedish students: putting the stress on the last syllable (fig.4).

IV. CONCLUSIONS

We have argued for a high-quality CALL infrastructure which should:

1) provide access to diagnostic tools, inclusive tools for automated self-assessment;
2) promote communicative and cultural competence by means of authentic materials, presented by means of multimedia;
3) utilize linguistic resources and language technology tools on all level of language structure;
4) be open and flexible (give the tutors and the learners possibility to create and store training materials on their own);
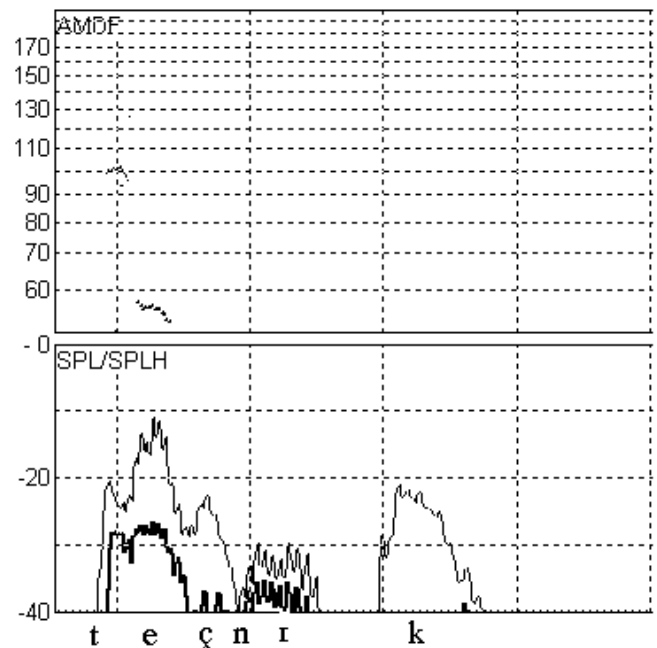

Fig.3. The correctly stressed German noun *Technik* (stress on the first syllable). SPL=Sound Pressure Level The upper part of the picture shows the voice fundamental frequency contour.
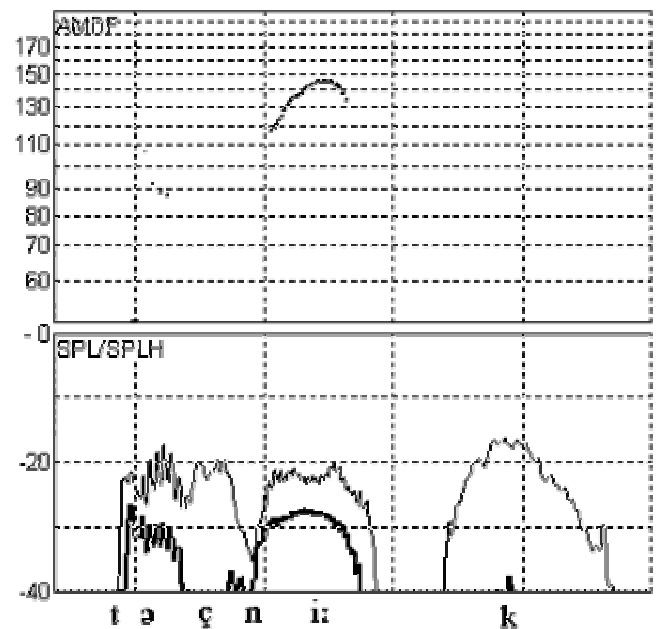

Fig. 4. The noun *Technik* pronounced with stress on the last syllable. SPL=Sound Pressure Level. The upper part of the picture shows transitions in the voice fundamental frequency contour.

5) be continuously evaluated in interaction between learners, teachers, and system developers.

The evaluation process should cover the following aspects (we adapt the definitions formulated in [25]):

1) Feasibility evaluation: concerns the question whether a particular approach has any potential for success;
2) Internal evaluation: tests on a continual basis whether the components of a prototype, or pre-release system interact as intended;
3) Declarative evaluation: deals with the coverage of the system (the vocabulary, the range of syntactic constructions, the different domains and topics);
4) Operational evaluation and usability evaluation: includes such questions as software and hardware compability, the costs of installing the system, naturalness of navigation, and user friendliness in general.:

As a final conclusion, we want to stress the fact that a CALL system should – even if it displays a very high quality – not be understood as a tool that would make the language teacher superfluous. It should be aimed at replacing the teacher in tasks that require searching in large databases, choosing audiovisual materials, simulating authentic environments, or confronting the learner with different language variants and different speakers. It should also facilitate evaluation and reduce stress factor for the learner (for example, one feels no doubt more comfortable when training pronunciation against a computer than in front of the teacher and more successful class mates). The human teacher is, though, not possible to replace by a computer, when it comes to explanation of subtle semantic, pragmatic and sociolinguistic nuances, as politeness rules, social conventions, or expressions of feelings.

REFERENCES

[1] P. Hubbard, "A Methodological Framework for CALL Courseware Development". In Pennington, M.C. and Stevens, V. (eds.): *Computers in Applied Linguistics: An International Perspective*. Clevedon: Multilingual Matters, 1992, pp. 39-65.
[2] V. Hegelheimer and C. A. Chapelle, "Methodological issues in research on learner-computer interaction in CALL", Language Learning and Technology 4, 2000, pp. 41-59.
[3] J. Egbert and E. Hanson-Smith (eds.) *CALL Environments: Research, Practice, and Critical Issues*. Alexandria, USA: Teachers of English to Speakers of Other Languages, Inc., 1999.
[4] Duber, J. (2002, November) Computer Assisted Language Learning @ CHORUS. Available at http://www-writing.berkeley.edu.chorus.call
[5] EUROCALL, homepage. (2002, August) http://www.eurocall.org/index.html
[6] K. Ahmad, *Computers, language learning and learning teaching.* Cambridge: University Press, 1985.
[7] Common European Framework of Reference for Languages (October, 2002). Available at http://www.coe.int/T/E/Cultural_Co-operation/education/Languages/

[8] J. Trim, *Gemeinsamer europäischer Referenznehmen fuer Sprachen: lernen, lehreh, beurteilen. Niveau A1, A2, B1, B2, C1, C2.* Berlin/Muenchen/Wien/Zuerich/NewYork: Langenscheidt, 2001.
[9] G. Hollenstein, "Ich lerne meine Sprachen kennen, oder? Von Referenzrahmen, Sprachenportfolio und neuen Dimensionen". *Ausblicke 16. Zeitschrift fuer österreichische Kultur und Sprache.* Zentrum fuer Österreichstudien, Jg.8, H.2. pp. 4-8, November 2002.
[10] DIALANG 2002: Version 0.8.0 (October, 2002). Available at http://www.jyv.fi/DIALANG/
[11] R. Bannert, *På väg mot svenskt uttal.* Lund: Studentlitteratur, 1990.
[12] I. Holst, *Datorstödd språkinlärning – ett kommunikativt orienterat prototypsystem.* B.A: Thesis, University of Skövde, Sweden, Dept. of Languages, 2002.
[13] C. J. Dodson, "The independent evaluator's report". *Schools Council Committee for Wales*, 1978.
[14] H. H. Stern, "Analysis and experience as variables in second language pedagogy". In B. Harley, P.Allen, J. Cummins and M. Swain (eds.), *The Development of Second Language Proficiency*, Cambridge: Cambridge University Press, 1990, pp. 93-109.I. Brown, H.D. (1994). Principles of Language Learning and teaching. London: Prentice Hall.
[15] B. Harley, P.Allen, J. Cummins and M. Swain "The nature of language proficiency". ". In B. Harley, P.Allen, J. Cummins and M. Swain (eds.), *The Development of Second Language Proficiency*, Cambridge: Cambridge University Press, 1990, pp. 93-109
[16] C. Brumfit, *Communicative Methodology in Language Teaching: The Roles of Fluency and Accuracy*. Cambridge University Press, 1984.
[17] M, Warschauer , H, Shetzer. and C. Meloni, *Internet for English Teaching,* Alexandria, USA: Teachers of English to Speakers of Other Languages, Inc, 2000.
[18] D, Healey. "Communicative Skill-Building Tasks in CALL Environments". In Egbert, J., and Hanson-Smith, E. (eds.): *CALL Environments: Research, Practice, and Critical Issues*. Alexandria, USA: Teachers of English to Speakers of Other Languages, Inc., 1999.
[19] G.A. Miller, (ed.), *WordNet: An On-Line Lexical Database*. Volume 3(4) of the International Journal of Lexicography. Oxford University Press., 1990.
[20] G. A. Miller, "WordNet: An on-line lexical database *"Communications of ACM*, 38(11), 1985
[21] P. Vossen, P. EuroWordNet. *A Multilingual Database with Lexical Semantic Networks*. Dordrecht: Kluwer Academic Publishers, 1998.
[22] A. Montoyo and M. Palomar, "WordNet Enrichment with Classification Systems". In *Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations* at NAACL 2001, pp. 101-106.
[23] B. Gawronska and B. Erlendsson, "Reducing Ambiguity by Cross-Category Connections". *WordNet and Other Lexical Resources: Applications, Extensions and Customizations. NAACL 2001 Workshop*, Pittsburgh, 3-4 June 2001, pp. 189-192.
[24] W. Wahlster, "Mobile Speech-to-Speech Translation of Spontaneous Dialogs: An Overview of the Final Verbmobil System". In Wahlster, W. (ed.): *Verbmobil: Foundations of Speech-to-Speech Translation*, Springer-Verlag, Berlin, 2000, pp. 3-21.
[25] J. White, "Contemplating Automatic MT Evaluation". In J. White (ed.): *Envisioning Machine Translation in The Information Future, Lecture Notes in Artificial Intelligence* 1934, Berlin/Heidelberg/New York: Springer, October 200, pp. 100-109, 0.