Language Processing and Intelligent Computer-Assisted Language Learning

Abstract:

This is the module "Sprachtechnologie für ICALL" from the MiLCA[1]-project funded through the bmb+f-initiative 'Neue Medien in der Bildung[2]'. The module contains information about the use of Natural Language Processing (NLP) in Computer-Assisted Language Learning (CALL). The *XHTML-pages* are available at http://www.cogsci.uos.de/~vreuer/milca/ICALL.xhtml[3] and use *Character Encoding UTF-8*. The pdf-version of this module can be found at http://www.cogsci.uos.de/~vreuer/milca/milca_icall.pdf[4]. The date of this version: 05.06.2004;

written by Veit Reuer (vreuer@uos.de)

1 http://milca.sfs.uni-tuebingen.de

2 http://www.medien-bildung.net

3 ICALL.xhtml

4 milca_icall.pdf

Table of Content

1. Introduction and Motivation

Abstract:

In the introduction the question of what the advantages of computers in language learning and teaching are, is answered and the content of the following chapters is described briefly. Learning Goals are the fundamental advantages of using computers in a learning scenario.

1.1. Introduction and Motivation

Using computers to aid the teaching of languages has been done probably since the late 1960ies in a number of institutions especially at universities. What was the driving force behind this? Several advantages can be noted theoretically for the use of computers in this domain and for the use of computers for learning in general. It is important to note that all of these aspects must be seen from the learner's perspective. Why the area of computer-assisted language learning (CALL) might be interesting from the perspective of computational linguistics (CL) will shown later.

Figure: *Immediate feedback*

One of the main reasons for using computers in a learning scenario is the possibility for immediate feedback to learner input. In a standard class room setting this can only be achieved by the teacher as opposed to doing exercises in a workbook. Using the computer, every learner has the possibility to receive immediate feedback.

Figure: *Different types of interaction*

Not only immediate feedback is possible but also different types of interaction with the computer. The learner may not only type and read, but look at a video, point at something with the mouse, or speak into a microphone.

Figure: *Time and place*

Additionally the use of computers is usually not restricted by time and place, whereas face to face lessons are always limited in this sense. One consequence is that the pace of learning can be adjusted individually.

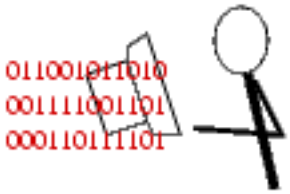

Figure: *Information in one box*

A fourth reason is the availability of various sources of information in one box. When the learner tries to do an exercise s/he can consult e.g. an electronic dictionary, the program's help-files, or the GWWW. The information can also be conveyed through different media (graphics, audio ...). Everything is coded digitally in one box.

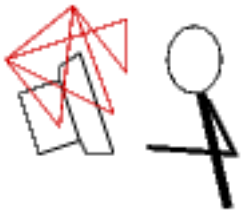

Figure: *Hypertext*

Finally the linear structure of printed material is lifted and material can be presented in a multi-level design, i.e. hypertext. Again this allows for an individual learning strategy according to the desired content and for a certain amount of knowledge discovery.

These advantages are the ones which concern the language learner independent of issues of computational linguistics (GCL). However from a different perspective, there are some reasons why methods of computational linguistics can be applied to GCALL, which is then called intelligent computer-assisted language learning (GICALL). One main advantage is, that the language domain in language learning is usually limited. The idea is that at least in beginners to medium language classes, language usage is restricted enough to handle it with the current state-of-the-art language technology.

Additionally to the limitation of the language, the world-knowledge domain can also be easily restricted with the design of the exercise. As with the language a limitation in world knowledge increases the chances that the system will perform as expected by the learner.

Finally the field of learner language can be seen as a challenge to CL. A lot of commercial CALL programs lack a lot of functionality and questionable types of exercises, as will be seen. Language technology might be able to improve the functionality and provide more and suitable information about the language either entered by the learner or contained in the program.

Since much of the research in language technology is concerned with the analysis of language (vs. the generation), this might also be applied in ICALL-systems. Analysing language input

which is probably in some way erroneous makes an interesting aspect for error recognition applications. For these reasons it might be interesting to choose CALL as a field of application.

1.2. Content

This page presents an overview of this course on ICALL as the following topics will be covered. The items in the list are ordered according to the planned schedule and not according to the table of contents.

1. In the first part the view of the learner is presented combined with the views taken in current didactics on computers and language teaching. It will be shown, that the dramatic increase in computing and communication power has led to an increased usage of computers in language teaching besides other applications. New types of exercises seem to have been established in language teaching with the aid of computers which enable the learner to access a foreign language with (possibly) new methods.

   Also the trends in language didactics seem to go along well with the development of information technology (IT): For example one of the goals is to allow the learner to use so called explorative learning. With the help of the WWW and some well structured exercises, this can be achieved in a straight forward way.

2. Then a short part about the history of CALL and more important ICALL is presented. This part should enable the reader to put the following chapters into a "developmental" context. The history demonstrates how closely interwoven the development of computer technology and of the usage of computers in language classes is. Some authors have even asked, whether the development of teaching might be driven by technology. More specifically the history of ICALL can give an impression of where researchers have seen and still see interesting areas to apply methods of computational linguistics to improve CALL-software or to implement completely new types of programs. First and foremost this is of course applying natural language analysis to learner input, but also other options are available as will be seen in chapter 3.

3. Following this, two (historical) example systems are presented, which show the main direction of research development in the last decades of the last century, namely to implement dialog oriented language tutoring systems. An important aspect here is (of course) the difficulties the developers had in developing the systems in the area of computational linguistics and, less important, in other fields. In both cases almost complete dialog-systems were implemented in order to be able to analyse learner input and to generate appropriate output. These systems therefore allowed the learner to (almost) freely produce language in order to enhance her/her so called communicative competence.

4. In order to achieve a further insight into the area of actual ICALL-systems, a categorization scheme can be employed. This is especially important with the language learner/teacher in mind, because a program should be used at the right level at the right stage of the teaching process. Furthermore especially ICALL-systems, which are usually only prototypes and not complete self-learn-programs, are developed to be used in certain learning

situations.

5. One important aspect about developing a program is an evaluation with respect to the intended usage. There are several aspects, which must be considered when evaluating a program for language learning. A certain evaluation scheme must be changed additionally, if ICALL-systems instead of traditional CALL-programs are to be evaluated. Considering an evaluation scheme also helps in the development of a new ICALL-system. Since the technical aspects of a project are usually in the foreground, an evaluation scheme might help also to keep the learner and his needs in mind.

6. The chapter on functionality tries to show the variety of places, where an integration of CL-techniques and linguistic resources can and has been tried. This ranges from superficially simple orthographic checking to syntax error recognition to linguistic databases. Basically two main areas can be distinguished: Methods (or tools) and resources. "Methods" in this context means applications, which do either analysis or generation of language. Resources however are not used by an application per se, but are usually data bases with linguistic information, which can be retrieved for presentation to the learner.

7. The following chapter deals with error analysis, and more specifically with the recognition of syntactic errors in learner language. This chapter is taken out of the list of error analysis subchapters, because in the field of CL this aspect has received the most attention. In general one can say, that there are two ends on how to recognize syntactic errors with NLP-tools. Either a parser is developed, in which the parsing algorithm is capable of identifying errors or the linguistic information, i.e. the grammar and/or the lexicon contain information about possible errors. Additionally a mixed approach can be chosen in order to exploit the differences in the predictability of errors.

8. Following the treatment of syntactic errors a look is taken at the recognition of errors in other linguistic levels, such as phonology, morphology and semantics. These fields are dealt with together, because only little reseach has been carried out here. Nevertheless these fields are important areas in language teaching and therefore should be mentioned.

9. The topic of this part is the construction of intelligent tutoring systems ([G]ITS). As the name suggests these systems are designed to tutor a user with respect to certain learning material. Here the chapter tries to abtract away from the content and offers insight into the advantages and difficulties related to ITSs in general. There are several modules which are suggested to manage the interpretation of the learner input, the inference with respect to the internal model, the presentation of pedagogically sound follow-up actions etc.

10. The last item of the course deals with learner modelling in ICALL-systems. Here the inferencing is not in the foreground but the general possibilities for adaptation of the sy-

stem to the language learner. There are various ways of adaptation with respect to the content, to feedback, to help-messages etc. On the one hand these adaptive measures should adapt the system to the user and on the other hand this should enable the system to provide the right learning material and exercises at the right level of progression, i.e. challenge the learner.

1.3. Exercises Introduction

Homework in the summer course 2003

Have a look at the WWW-language-pages of the BBC[5], and write down your opinion about selected exercises. Make a scheme about the areas of your evaluation before actually trying some exercises.

5 http://www.bbc.co.uk/education/languages/

2. Overview

Abstract:

The following parts contain general remarks about the way language learning should be taught based on didactics research and how this is connected to using computers in language learning. The second part is about the ICALL-research, which has taken place over the years. There have been various important initiatives to develop *intelligent* language learning systems, but the outcome of these projects rarely managed to surface as a commercial product. Two example ICALL-systems are presented in more detail. Learning goals are a general "feel" for the issues involved in the didactics of CALL and a glimpse into the history of ICALL.

2.1. Didactics I

Rüschoff, Wolff, 1999 use a model of foreign language learning/teaching which incorporates three different goals: the development of *communicative competence*, of *language awareness*, and of *language learning awareness*. Language proficiency therefore includes not only the use of a language but also the knowledge about a language and knowledge about learning a language. This demonstrates that not only the goal of using a language is important, but also the path to achieving this goal, i.e. not only a language is learned but also learning a language is learned. Connected with these three areas are very different classroom activities. A similar view can be found e.g. in Ritter, 1995, but there are also widely disagreeing ideas about the goals of foreign language teaching, e.g. Storch, 1999. It has to be noted that these differing approaches tend to neglect the use of computers for language learning and therefore will not be considered here.

Basic building blocks for language teaching are considered the following (according to Rüschoff, Wolff, 1999):

- ***Process orientation***: All methods for acquiring communicative competence should consider the (psycholinguistic) processes involved in the production of language but also in the understanding of language. The same is true for processes involved in language awareness and language learning awareness. As an example one might consider the use of as many "cognitive channels" as possible when presenting language material. Obviously this supports the memorization of knowledge, i.e. the process of acquiring a foreign language is more "intense". Another example is the consideration of acquisition steps. It is well known, that the acquisition of language goes through certain levels independent of the content of the teaching, which are (at least in the early stages) clearly identifyable. The teaching therefore should be pointing towards the support of the actual stage the learners are in.
- ***"Action orientation"/Project orientation***: The action in a language education scenario should always be directed at the use of the foreign language in communicative situations. The action in lessons based on acquiring communicative competence used to be limited to simulating the action in role-play, games etc. Using new technologies, i.e. the internet, this can be changed to having some form of a *true*, albeit limited communication situation. Projects are used to enlarge the goals of student action to higher goals than for example doing a certain exercise. Over a fixed period of time students learning a foreign language are working together on projects about a theme connected with the foreign language. This aspect is closely connected with social learning and autonomy of the learner.
- ***Authenticity***: This can be understood in two ways: Authenticity of the classroom action and authenticity of the material used in the classroom. Both are highly desirable. Of course both have to be used in an appropriately supported way. The material presented to

the learner should on the one hand be a little demanding, and on the other hand the learner must be able to cope with the material. Excessive demands must obviously be avoided. With respect to the use of computers in language teaching the rapid development of the WWW has made it very easy to obtain a lot of authentic data. The teacher does not need to collect and provide the learning material himself anymore, but can delegate this task to the learners themselves. The "authentic" communication with the computer as a tool is limited to a small set of communication types for obvious reasons.

- ***Social learning***: Working together in small groups has various advantages when learning a foreign language. "Nur in der Kooperation mit anderen kann es zu einer Angleichung der subjektiven Wissenskonstrukte kommen, können die individuell formulierten Hypothesen getestet und die subjektiven Konzepte aufeinander bezogen werden." (Only in cooperation with others can subjective knowledge be adapted, can individual hypotheses be tested and can the subjective concepts be related to each other.) Rüschoff, Wolff, 1999, page 63. However it seems questionable, if the use of computers in language learning supports some kind of social learning in any special way. The mentioned project work, which makes some sort of social interaction necessary, is not dependent on the use of computers. Social learning not related to language learning therefore might only occur, when learners are working together in front of the computer.
- ***Autonomy***: Learner autonomy should possibly not be understood as the learner using a computer autonomously without the guidance of a teacher. According to Rüschoff, Wolff, 1999 this can lead to a dangerous development, since learners should not be "left on their own in front of a computer". Learner autonomy should be understood as enabling learners, i.e. students in schools, to act and decide by themselves.

Following these didactic elements the main question for ICALL can be derived: How can these concepts be supported by CL-technology in CALL-programs (if at all)? On a more abstract level the goals of language teaching must be considered: How can the acquisition of communicative competence, language awareness and language learning awareness be supported by computational linguistics' tools and methods?

2.2. Language Didactics and Computer Programs

Looking at different types of computer programs, different aspects of language didactics come into play. In Rüschoff, Wolff, 1999 it is strongly argued for the integration of general modern language didactics like learner-autonomy and project-orientation and the use of information-technology. The authors therefore encourage the use of computers not as tutoring systems but more importantly as explorative devices. The learner is invited to explore the language he intends to learn. With the wealth of material in a foreign language now available in the computer via the internet, the teacher can give hints about what to look for and then the learner autonomously discovers the properties of the language. This type of usage is fundamentally different from the use of a computer as a tutoring system. Taking a look at what different categories of computer programs can offer things may be seen in a different light.

The main advantage of computers over pen and paper when solving exercises in so called tutoring systems ("Functionality", S. 36) is the immediate feedback, that the learner can get from the computer, as mentioned earlier. Grüner, Hassert, 2000 caution against the term "interaction", i.e. interactive exercises, and try to distinguish "communication" and "interaction" with the computer. In their opinion the term "interaction" is not plausible, as describing what the computer is capable of. They therefore suggest merely to use "reaction".

Even though *tutoring systems* are very restricted concerning the learning material and the learning progression and usually can not be adapted to personal preferences, there are certain circumstances, when tutoring systems might be of use. Usually tutoring programs are considered to support especially the learning of grammatical structures and lexical elements. The addition of multimedia-material might support the learning process, albeit this has not lead to new types of exercises with respect to the usage of language.

Teachers can use *authoring programs* to develop their own tutoring systems. This allows the teacher to develop exercises especially suited to the needs of his/her students and the learning-situation they are in. Unfortunately developing such exercises takes up a lot of time. Estimates are that the ratio of developing a learning unit to working on the exercises is around 10:1.

Using *computers as resources* for language learning involves mainly electronic dictionaries. Fast and combined access to even multimodal information is one of the advantages compared to printed editions. Some dictionaries include the possibility to extend the dictionary with user generated items. Encyclopedia also belong to this category of electronic resources which present monolingual, "authentic" information to the user. Resources which need the internet, e.g. databases on servers accessible through the internet, are listed below under "communication possibilities".

Programs such as text processing systems or concordancers can be used in language learning as *tools* "to actively model the difficult process of acquiring a second language" (Rüschoff, Wolff, 1999 p. 73). This plays an important role in designing process- and action-oriented

language teaching. The tools can be used by the learner themselves to discover properties of the foreign language.

Using todays *communication possibilities* has at least three main aspects: 1. The wealth of authentic material such as texts, graphics and video offered by the WWW partly in databases; 2. the possibilities for electronic communication with native speakers of the foreign language with E-Mail, Chat etc. and 3. the possibilities of developing distance-learning concepts. All three components (communicative competence, language awareness and language learning awareness) can be addressed with the use of communication technology.

Some more literature about using computers in language teaching with a focus on German as a second language: Fechner (Hrsg.), 1994, Bayerlein, 1996, Eidecker, 1996, Goettmann, 1996, Hess, 1998, Rösler, 1998, Wolff, 1998, or works by M. Warschauer[6] e.g. Warschauer, 1996. There is -of course- a lot more in the journals mentioned in the "Resources", S. 59 .

6 http://www.gse.uci.edu/markw/papers.html

2.3. Didactics II: Learning Skills

Computers have been part of the tools to support language learning since the late 60ies of the last century. Considering the development of computer technology of that time, most of the programs of course were simple text-based systems, which mainly allowed for some kind of simple exercise-solving, namely multiple-choice or gap-filling. This means that listening and speaking were not trained and only little writing. With the advent of the capabilities of multi-media, computers have become supportive to all skills in language learning: Listening, reading, writing and speaking.

- ***Listening*** can be trained by playing sound-files, if the computer is able to produce sound. Especially the understanding of texts can be supported. After listening to a text being read, the learner may be asked to work on some exercises about the text. Two major advantages can be noted over the tape-recorder. Firstly the computer can give immediate feedback to the exercises even with reference to the text/reading-position and secondly the sound-files can be replayed with very high precision.

  Also oral dialogues can be heard. In this case transcripts can accompany the sound-output to further enhance the understanding. In contrast to the classroom situation this lets the learner train his/her capabilities for understanding language produced at a "normal" speed, since fellow learners and/or the teacher tend to produce language on a level a lot more tuned towards the proficiency level of the hearer.

- ***Reading*** can be trained with various types of exercises following the presentation of texts. But there are some serious, more technical difficulties with todays sreens. Longer texts can only be presented partially which hampers the overview. Some texts (e.g. hypertexts) do not allow adding personal notes (or only with difficulties). Reading of longer texts may be difficult because of poor refresh-rates of some monitors. Additionally there is still a kind of "usage-barrier", which prevents learners to use electronic equipment such as computers and e-books as a replacement for printed material.

- ***Writing***-skills can be practiced using a computer with some advantages. Texts produced with some kind of word-processor or editor are 'clean' and easily readable. The understanding of texts can be enhanced with graphical presentations added to the text. Most word-processor have build-in orthographic checking, which can be used to make the texts almost free of orthographic errors. Albeit checking for morphosyntactic errors is unreliable and semantic errors can not yet be handled. Most word-processors also include a thesaurus. This could be used by language learners to change the choice of words in their text. However this requires some knowledge about the use of words and phrases in the target-language. So far, there is very little support for learning writing except for the field of dictation.

- ***Speaking*** into a computer is still a minor feature. The voice of a learner can be recorded

and in most cases has to be evaluated "manually" by the user himself. Few programs include an analysis-module for evaluating the soundinput. This evaluation usually amounts to the presentation of percentage-figures of the difference between the learner recording and the prestored sound pattern. In newer programs, spectrograms or the like are presented to the learner. If one knows how to interpret this data, it should provide some concrete help instead of a general match/mismatch-feedback.

As in the previous page the goal of enhancing CALL-programs with the help of CL one should keep these four basic skills in mind in combination with the afore mentioned learning types.

2.4. History of ICALL

Most of this part was taken from Levy, 1997. As opposed to the previous chapters, from here on the focus is on *Intelligent* Computer-Assisted Language Learning (ICALL). Even though there may also be references to traditional CALL here the part about ICALL starts.

Prototypical ICALL-programs were developed as early as the late 70s, e.g. Weischedel, Voge, James, 1978. These programs ran on large mainframe-computers in the computing-labs of the universities. One of the large projects of the last century was the ATHENA Language Learning Project (ALLP) at the Massachusetts Institute of Technology (MIT) which started in 1983 (Murray, 1995). In the project the method of communicative language learning was to be applied to a computer program. Three main technologies were used to develop the system: Natural language processing, speech processing and interactive video. One part of the material presented to the learner was a story about an amnesiac Columbian scientist. Interactive video was used to present some part of the story and then the learner had to complete the information through interactive dialogs with characters from the plot. Another task was the communication with a Poltergeist, who could be orderd to clean or mess up a room according to the instructions given by the learner. This included not only natural language parsing, but also discourse processing (focus) and a model of the world, which could be altered according to given limits. It has to be mentioned, that the ALLP was only in certain parts used with language learners for an evaluation. More information about this specific system can be found in the chapter "Example System I: ALLP", S. 20 .

Another project, which was developed a little later, was FLUENT-1 and FLUENT-2 by Hamburger, e.g. Hamburger, 1994 and Hamburger, 1995. Here a similar approach was chosen, in that the learner was to clean up a kitchen, in which a graphical hand moved according to the learner's input. The input could be done in natural language or via moving the graphicals. The concept behind this system was to integrate the visual approach with the language approach to 'immerge' the learner into the situation without recursion to grammar, vocabulary or mother tongue. More information about this specific system can be found in the chapter "Example System II: FLUENT", S. 23 .

According to Levy, 1997 one problem of these systems was, that they were only tested in artificial situations, where the developers tried to forsee the reactions of the students. More recent developments try to incorporate more "real" data to make the systems capable of handling data, which occurs in real language learning situations e.g. the ICICLE[7]-project. Even though the mentioned programs ALLP and FLUENT were largely prototypical systems I do not know, why e.g. the mentioned commercial program "Herr Kommissar" (deSmedt, 1995) is not available anymore, even though it seems to have been used successfully.

7 http://www.eecis.udel.edu/research/icicle/

As far as I know there is not a single "intelligent" language learning program on the market available in Germany right now. The only true advancement which can be identified is that there seems to be a trend to incorporate more multimedia capabilities into the programs which supports the multi channel requirement. The disregard of methods of computational linguistics might also be due to the increased usage of the internet to stimulate the general use of computers in language learning. Some programs now include sections, where single words can be spoken into a microphone and the program tries to evaluate the correctness of the pronunciation. However this should probably not be called an "intelligent" program, since these programs do not include any mechanism to identify words of a language as e.g. speech recognition software does.

It seems worth to mention in this chapter, that resources and linguistic tools have become a valuable part of the concepts in CALL and also ICALL. In the 1980ies computers were not fast enough and/or did not have enough memory to store and provide language resources like lexica and corpora. This has changed dramatically in the last few years and is likely to continue. There have been a number of concepts being developed to use resources such as WordNet[8] or CELEX[9] as well as corpora for language teaching/learning purposes.

In recent years several programs have been developed in research projects which on the one hand include techniques from CL but on the other hand are actually used in language teaching, such as the E-Tutor[10] and Glosser[11]. These projects demonstrate that computers now have matured enough to be able to run complex NLP-applications.

An important external link for more information is The history of CALL[12], even though it contains a lot of information about the general development of IT and ends around the year 2000. All important research initiatives during the last decades are mentioned and set into context with language teaching and IT-development.

In the next sections two systems are presented in more detail to give an idea of the advantages and diffculties which show up in the development and usage of ICALL-systems.

8 http://www.cogsci.princeton.edu/~wn/ - Last Visit: 6.2004

9 http://wave.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC96L14 - Last Visit: 6.2004

10 http://www.e-tutor.org - Last Visit: 6.2004

11 http://odur.let.rug.nl/~glosser/ - Last Visit: 6.2004

12 http://www.history-of-call.org - Last Visit: 6.2004

2.5. Example System I: ALLP

In this large-scale project started in 1983 an intelligent language learning program was developed. Three main new technologies were introduced via this program into the classroom. 1. Natural Language Processing, 2. Speech Processing and 3. Interactive Video. The information in this section is mainly taken from Murray, 1995 and Felshin, 1995.

Three major exercise types were implemented in the ALLP-framework. The first one was a conversation simulation with a poltergeist, which could be directed to clean up or mess up a room according to the learner's input. This was called *LINGO*. In this microworld relatively complex dialogs were made possible through the use of a complete dialog-simulator, described below. The poltergeist was not only able to make suggestions of what to do next, but also had a concept of things being in focus. Therefore e.g. the last mentioned item could be referred to with a pronoun. Also the character of the poltergeist could be changed with the use of language either being of help to the student or hindering the actions of the learner.

The second exercise was called *No recuerdo*. This consisted of an interactive video story about the adventures of an amnesiac Columbian scientist. Here the learner played the role of a journalist who had the task to "uncover the truth": the learner was asked to interview various actors in the plot. The responses from the characters were either recorded and played from the video disk or they were created by some NLP-module. For both scenarios - the poltergeist and the scientist story - the system has not been tested with language learners because of hardware limitations, even though demo-versions were produced.

The last main type was an *Intelligent Workbook*, which offered a series of interactive grammar exercises. A situation was presented and some English paraphrases were given for the learner to produce a correct Spanish utterance, which effectively meant a translation task. Here also some hardware limitations did not allow the system to be implemented fully. But according to the publications "real" learners were able to use this part of the program.

The NLP-System

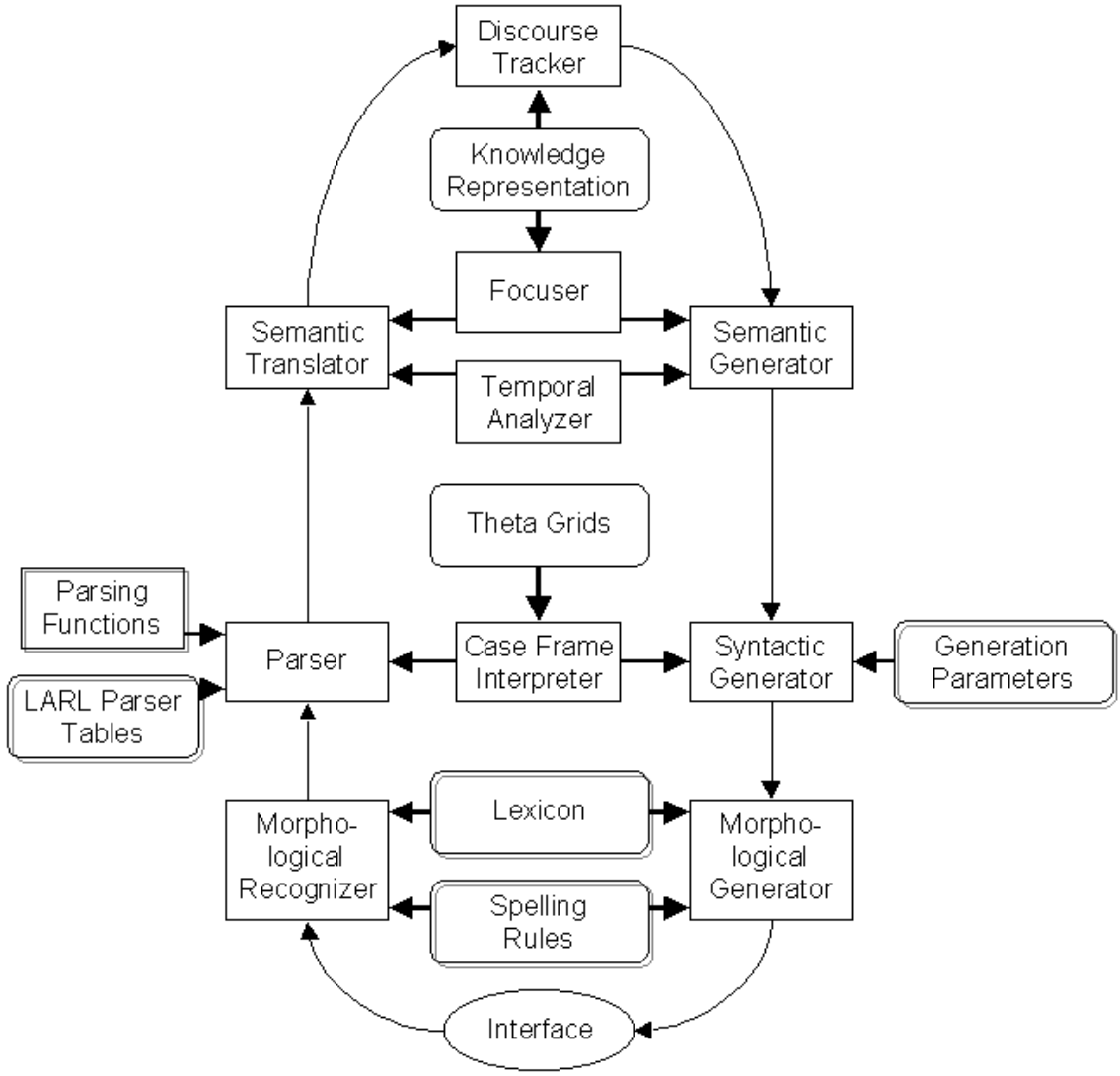

Figure: ***Design of the NLP-System in the ALLP***

In the above figure taken from Murray, 1995 one can clearly identify all the major parts necessary for a complete dialog system. Square boxes mark code whereas rounded boxes signify data. The shadowed boxes show the language dependence of some modules.

The system was developed right from the beginnig as a language independent system with clearly defined modules for different languages. Tests were done with Spanish, French, German, English, and Russian. Therefore the system relied heavily on a type of *interlingua*, which was used in the discourse module. Errors in learner's input were detected in two ways. Agreement errors were identified by feature relaxation whereas word ordering errors were recognized via additional rules in the grammar. The system was also able to identify very uncommon morphological or syntactic forms and to apply penalties to these even though they

were actually correct. More common interpretations of the input were then chosen even though this might mean to hypothesize an error.

The system contained a complete morphology generation subsystem, which was able to generate most surface word-forms from underlying stems and affixes. The syntax was modelled after the Government-and-Binding theory (Chomsky, 1981) using concepts such as S-, D-structure, and CF-structures. However the parser could only generate one analysis-tree at a time and was not being able to perform movements. In order to retrieve semantic interpretations of input the lexicon and the grammar made heavy use of thematic roles and the so called case frames. The following example shows the frame for the English verb "tell".

Example1: allplexikon

```
(:voice active
(:thematic-role agent
:syntax((:case subject :type dp :required-p t)))
(:thematic-role theme
:syntax((:case empty :type vmax :spec (or indicative infini-
tive))):required-p t)))
(:thematic-role destination
:synatx((:case indirect-object :type dp :required-p t))))
```

In this lexical entry one can identify some more information about the structure of the complements of tell (type vmax), which must be either indicative or infinitive.

The parser used in the system was a LALR-parser, which essentially followed the Tomita-algorithm (Tomita, 1986). The stack could be accessed and through this flexibility was achieved. Some semantic processing was done during "syntatic" parsing in order to eliminate unlikely paths as early as possible. The parser was written in Common Lisp and C.

The system included different grammars for recognition and generation. Since the recognition grammars were also partly responsible for error recogniton, the size of that grammar was about 10 time bigger than the generation grammar, according to Felshin.

The only references to the ALLP-project on the WWW I could find, are LINGO[13] and No Recuerdo[14] (last update 1997!).

13 http://www.improbable.com/personal/gorin/artwork/lingo.html - Last Visit: 6.2004

14 http://web.mit.edu/fll/www/projects/NoRecuerdo.html - Last Visit: 6.2004

2.6. Example System II: FLUENT

The FLUENT-program mainly consisted of a microworld, in which the student could act and the "tutor", i.e. a language generation module would respond to these actions. In a second type of exercise the tutor could command the student to do things in the microworld and learner was expected to act accordingly. The system therefore resembled the ALLP-System with the "Poltergeist"-scenario in a certain sense. Most of this text was taken from Hamburger, 1994 and Hamburger, 1995. The microworld consisted of a kitchen scene with a sink, a stove, and several cabinets as can be seen below. The aim of the system FLUENT-1 was to provide a connection between graphical action (in a microworld) and language output of the system. This was done in order to immerse the learner into a scenario, in which only the target language was visible and available.

The following figure is a scan from Hamburger, 1994. It shows a graphical hand (above the sink) which could be moved by the learner to do various things in the kitchen such as filling a cup or putting a pot on the stove.

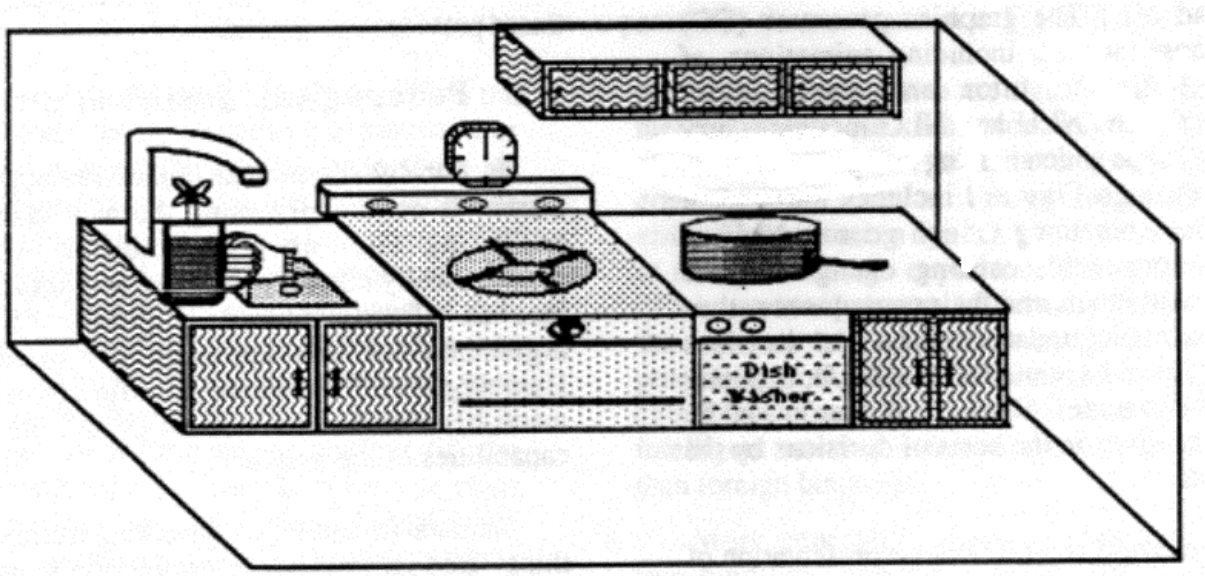

Figure: *Example scene from FLUENT-I*

The NLP-component of FLUENT-II is the same as the one used in the ALLP-prject. According to Hamburger, 1994, it was specifically adapted to be used in the FLUENT-concept. Therefore not only a language understanding module but also a generation module was included.

One important aspect is the recognition of different views on a single action or scene in a computer-program. The picking up of the pot in order to move it can be viewed from very different perspectives, which are verbalized differently. The modelling of these views allows for a great variety of verbalizations of the same facts. Additionally Hamburger claims, that this achieves a tight integration, but not interdependence, of graphics and language. The actual

verbalization by the system depends on the state of the microworld in combination with the learner-model.

It seems however, that this system has not been evaluated in a true language learning setting. Therefore it is difficult to say whether the advantages of the system actually hold.

2.7. Exercises Overview

Homework in the summer course 2003

Special presentation: Develop a little timeline (about 20 items) with the general developments in Computer and Information Technology and present it to the course in the next meeting.

Special presentation: Develop a presentation on the main features of so called Dialog Systems with a focus on Computational Linguistics methods.

Clearcut questions

What are the four basic language learning skills?

What are the three goals of language teaching according to Rüschoff and Wolff 1999?

What are the possible advantages of using a computer for language learning and why?

What are the reasons of why there isn't any CL-technology in CALL-programs today?

Outreaching questions

What might be the advantages of the internet with respect to language learning?

What other reasons can you think of for using methods of CL in a CALL scenario?

How can especially methods of CL improve the four language skills mentioned above?

Is CALL technology-driven or usage-driven?

3. Categorization and Evaluation

Abstract:

In this chapter a categorization scheme for the various types of programs which can be used for language learning and teaching is suggested. This is (at first) done without regard of methods of CL and AI. The next chapter on Functionality will build on this structure. In the second part some aspects of program evaluation are explored. Both aspects are important as most systems are not complete language trainers but focus on certain aspects of language learning. This applies to ICALL systems even more so.

3.1. Categorization Introduction

Computer programs, which can be used for language learning, can be categorized in several ways. One possibility is the following:

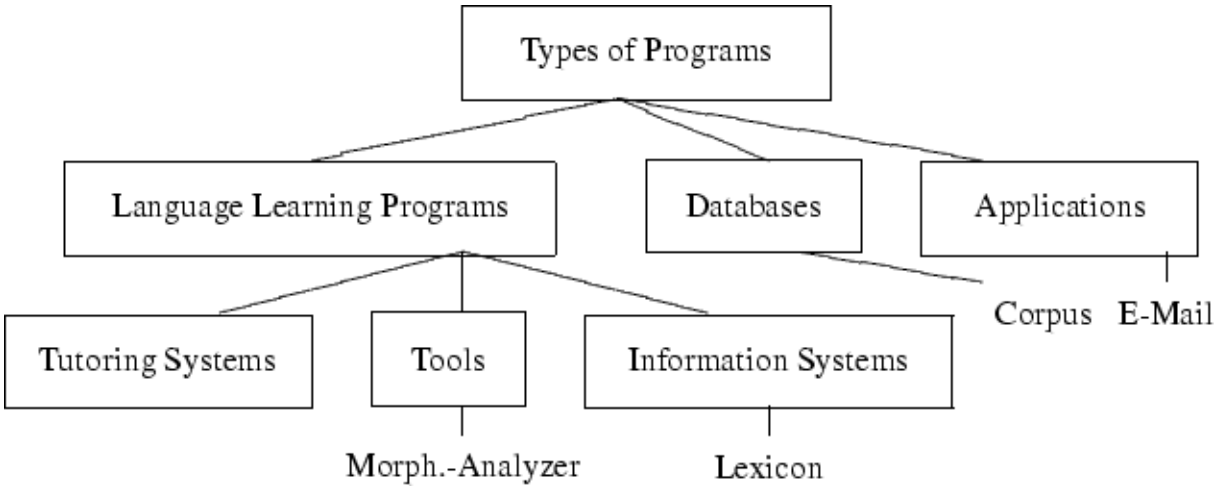

Figure: Categorization

Wolff, 1998 has a different cassification. On the first level he distinguishes between tutoring systems, tools and communication applications. This shows e.g. how important he thinks telecommunication nowadays is for language learning and teaching. Other characteristics of computer applications can be used too. Ritter, 1995 also uses three categories, namely tutoring programs, games and simulations as well as tools.

Another classification (Thome, 2000) distinguishes 8 types of programs: Learning- and training-programs, intelligent tutoring systems, simulations and experimental environments, multimedia "Erlebnisumgebungen", thematic databases, environments for communication and cooperation, tools for writing, designing etc., as well as programming environments.

This chapter will only introduce programs, which belong the category of "language learning programs" (see figure Grafik "Categorization"), i.e. which were designed specifically to teach or to train languages. Other programs such as textual databases or applications like word-processors can also be used to teach or train a language. There are some language-teaching specialists who prefer the latter types of systems, because they interact nicely with modern types of didactics as mentioned in the chapter about didactic issues. One important didactic theme is that of "explorative learning". The learner is offered a lot of authentic material which he is able to explore and able to use for building his own hypotheses about the language. These kinds of programs are not at issue here, even though some of them might even be enhanced using CL-technology.

The next chapter about "Functionality", S. 36 will go into more detail in this matter.

3.2. Categorization Tutoring Systems

Tutoring systems can be defined as programs which try to teach various issues about language. These could also be called "classic" language learning programs. Rüschoff, Wolff, 1999 define this class of programs as having the main task to give the learner to opportunity to train lexemes and grammatical structures of a foreign language.

Some characteristics are mentioned in the following.

In traditional CALL-systems there is a certain range of possible exercises. Multiple choice, gap-filling or sorting tasks are used in a lot of variants. This is primarily done to learn vocabulary. In some programs these exercises are also associated with certain grammatical phenomena, which are especially trained. Examples are the use of articles in German or AcI-constructions in English. These exercises often include some help-system, in case the learner does not know what to do either using the program or working on the exercise. At the end of a unit or at the end of the program often a test can be done which is different from the exercises in that the system responds only after the completion of the test and does not provide any help in between. Since computers nowadays can also be used to record sound, a new type of exercise emerged. Some systems have a function build in, which allows a recording to be analysed, i.e. matched to stored patterns. The learner can record his voice and the system then can tell the learner in how much of a percentage his recording diverges from the stored "correct" pattern.

Usually tutoring systems first provide some information either as text or as a soundfile. This information should be understood or learned by the user. Some help-functions are usually also provided. After reading or hearing a text an exercise follows based on the text. A "text" in the sense here can mean anything from a dialog or poem to a scientific text or any other coherent piece of language. Newer programs also include videos, which should be understood by the learner in order to work on subsequent exercises.

Multiple-Choice- and Fill-In-exercises and variants thereof are still the most used type. Even though nowadays multimedia and mouse enhance the old text-keyboard-based exercises, the general structure has not changed much. E.g. fill-in-exercises can now be realised in a drag-and-drop-manner. The addition of multimedia in these programs has changed the structure of the exercises, which the learner has to do, only very little. Even though he might now hear a text, he still has to show his understanding by doing multiple-choice questions.

As mentioned at the beginning students can learn at their on pace, i.e. there isn't a teacher who puts pressure on them to finish on time. Also a program can present the exercises many times (tireless) if some of the tasks were not solved correctly. This means, that the learner can be sure, that no topic will be forgotten.

Students can get immediate feedback. With the above mentioned exercises it is simple to im-

plement the code for answer checking and reacting to the learner's input. Some slightly more explorative exercises include text-reconstruction and sentence-switching. Nevertheless the programs usually provide some feedback for these exercises, too.

Another main characteristic feature of this type of program is that the learner is lead through the program by a tutor-like guidance-system. This guidance-system is usually based on a fixed path through the content, which might be altered be the completion or not-completion of previous exercises. However this does not mean, that the specific results from working on the exercises are used but rather that simply the learner is reminded of an earlier section if there are exercises not solved so far or answered wrongly.

3.3. Categorization Tools

Tools can be describes as applications, which tell the learner something objective about a language or rather a certain piece of language, i.e. tools can be considered online language analyzers. These programs help the learner to discover properties of the language he is learning. Rüschoff, Wolff, 1999 use the term "data-manipulating systems". The most simple form of such a tool is the search-and-replace function of a word-processor. With this tool the learner can search for certain patterns which might interest him. A different example tool is a concordancer, which shows certain word forms of a text or even a frequency analyser, which shows the frequency-distribution of a spoken piece of language to analysed by the learner.

Concordancers are programs which allow the user to search for patterns in large corpora with the help of (almost) [G]regular expressions. This can enhance the work with texts for students because it allows them to discover properties of the language themselves. Usually the data found in the corpora is pretty printed for easy overview. Thus not only lexical items can be extracted but also grammatical structures can partly be identified and analysed. This (of course) works best if the inflectional system of the language is not too rich as all possible wordforms have to be spelled out. Learners can then work on exercises regarding the content of texts and also regarding the linguistic features of texts.

One modern form of tool is build by publishing companies of textbooks who offer wordprocessors add-ons, which include databases with the vocabulary from their main language learning text books. Foreign texts which are loaded into the wordprocessor can then be checked against the vocabulary of the course. Unknown words are marked and can be explained by the teacher. A disadvantage is of course, that it uses only the database from one publisher. Additionally these add-ons can also identify prepositions, pronouns and other 'function words' to allow the teacher to create exercises.

One might also count translation systems as tools for language learning. Even though translation should be learned by students one can imagine situations in which translation tools might help the learner to support understanding. This is the case when the task is working with a text and not e.g. learning vocabulary.

Note that for the category "tools" immediatly some techniques used in CL come to mind to integrated in CALL applications. This has not been mentioned here as it will be the topic of the corresponding subchapter "Functionality and Tools", S. 39

3.4. Categorization Information Systems

Finally information systems are programs, which usually consist of a large database about a language, e.g. lexical databases or electronic dictionaries/grammars. Here no processing of the input takes place except for e.g. some orthographic checking if the word entered is not in the database. The usage in a classroom setting is therefore restricted to certain tasks as the programs cannot be adapted in a certain sense. Examples are the LEO Dictionary[15] and maybe Glosser[16]. The LEO Dictionary is at the base a German-English electronic dictionary interlinked with various other lexical databases. For example entries are crosslinked with the online Merriam Webster Dictionary and a inflection module producing the different word forms of a certain stem. Glosser presents texts with clickable words. When clicked different frames in a webbrowser show various information about the word such as morphological structure, translation and an example from another corpus.

Another online information system ist the Chemnitz Internet Grammar[17], which is a clickable grammar for English. Here the learner has to know what he is looking for. It is not for learning but rather for providing information on the structure of English.

The online version of the Merriam Webster Dictionary[18] is one of the few larger dictionary freely available on the WWW. The dictionary does not only provide information about the meaning of a word but also th other usual lexical information such as sound and grammar patterns.

In distinguishing tools from information systems one could say, that tools require some sort of effort from the learner or the teacher and the information received has to be interpreted. In an information system this has already been done and usually new data cannot be added.

15 http://dict.leo.org/

16 http://odur.let.rug.nl/~glosser/

17 http://www.tu-chemnitz.de/phil/InternetGrammar/

18 http://www.m-w.com

3.5. Evaluation

There are several online-sites which provide useful hints how to evaluate CALL-software. Here is a list of some of them:

- Evaluierungskriterien für multimediale Lernprogramme. Ein Raster für die Praxis[19] Schröder, 1996
- A Place to Start in Selecting Software[20] Healey, Johnson, 1997
- Evaluationskriterien für sprachliche Multimedia-Software[21] Wazel, 2000

These schemes were created to give a language teacher a check-list for evaluating CALL-programs. For the purpose of doing research on CALL-programs from a computational linguistics perspective these can also be used as guidelines. Ultimately a program using CL-technology should perform very well (and better) in a selected area. As is explained in more detail in the next section, programs can be looked at from two sides. One is the computational view and the other is the linguistic one. This could also be divided into a technical and a contentwise view. Considering this the evaluation then either focuses on how the program works, which types of exercises are presented, in which way the error recognition works etc. The focus may also be on what material is presented, how is it prepared and which linguistic areas are covered.

There is also the possibility to develop a specific questionaire for certain aspects of CALL-programs. In this case the questions have to be adapted to the aspect. Some aspects might include the help-system of a CALL-program or the possibilities for data-storage and learner-modelling.

The items of a checklist then have to be adapted also according to the available methods of CL. A question like "Is the program able to lead a speech-based dialog?" is then not valid because this kind of technology is not yet available.

The main evaluative questions with respect to the integration of methods of CL need to be directed towards the capabilities of the CL-application. Not just any CL-based application is usable in a language learning context. It must be clear to the learner whether some linguistic knowledge is required to interpret the output of the program or if the output can be understood very easily. The application can also be "disguised", in that it only supports the learner in the background. Eg. tools for corpus queries may not present the raw data but may be embedded in an environment.

A second area is the question of whether the integration makes sense with respect to new di-

19 http://www.sw2.euv-frankfurt-o.de/Publikationen/FsU/frame/

20 http://oregonstate.edu/~healeyd/cj_software_selection.html

21 http://www.iik.com/theorie/theoretisches/kriterien.html

dactic concepts of learning/teaching. A program, which still presents vocabulary exercises for drilling vocabulary lists is still didactically questionable, even if CL-methods are used for e.g. the generation of the exercises. When a concept is evaluated this has to be taken into consideration, too.

3.6. Exercises Categorization and Evaluation

Homework in the summer course 2003

Develop three different categorization schemes from different perspectives such as computer scientist, language teacher, learner etc.

Clearcut questions

Try to define the characteristics of information systems!

Do you agree with the overall categorization? If not, why?

What are three main areas to look out for while evaluating a program?

What is so difficult about articles in German or AcI-constructions in English?

Outreaching questions

Can you think of yet another categorization? What is the basis for your categorization?

How can CL-technology improve information systems?

Which categories of CL could be considered while evaluating a system?

How important is the right age and knowledge level of the user compared to other things like correctness of information in the system or program stability?

4. Functionality

Abstract:

A functional perspective on CALL-systems allows an assessment of the possibilities for CL integration. It seems clear that the ultimate goal of applying CL methods should be the improvement of functionality. In taking this perpective two main areas of application can be identified. Either the program may be improved without considering language in the first place or the "content" can be improved by using CL methods for the analysis of language data either from the learner or for the learner.

4.1. Functionality

In this part areas are pointed out in which computational linguistics methods could be used in a CALL scenario. This is one of the main aspects of applying CL-technology to CALL-systems: Functionality of programs should be improved with CL. The hypothesis is that methods of computational linguistics can support the learning of a second and even first language by using "intelligent" computerprograms.

Two main views can be taken. On the one hand one can start thinking about methods to enhance the computational aspects and on the other hand one can start thinking about the improved processing and presentation of content. An aspect of the first type could be advanced error recognition (see chapter "Error-Analysis", S. 43). An aspect of the second could be precise modelling of the morphosyntax of future tense or the integration of WordNet-like structures.

Probably there isn't a sharp division but a spectrum between these two views. Corpus linguistics is a case in question. On the one hand corpus-tools allow the learner to use large language data-bases and on the other hand specific items like collocations may be taught using corpora, see e.g. Erpenbeck, Koch, Kummer, Reuter, Tschorn, Wagner, 2001

If one accepts the categorization made in the previous chapter, then the question is: "Which functionality can be improved or newly integrated based on each of the categories?"

4.2. Functionality and Tutoring Systems

In traditional CALL-systems there is a certain range of possible exercises as mentioned above. These are restricted in the sense, that the pattern of solving the exercises is always very similar. In most cases the learner either selects an item from a list or the context in the exercise is limited in order to allow for only one sensible answer. This is the major disadvantage since the didactics call for free production of language to achieve communicative competence. Forcing the learner to use his personal knowledge of language has a much bigger impact on the learning process than choosing from a set of answers.

Even though most exercises are intended to train the use of language in communicative situations, the exercises with more or less predefined answers actually used are integrated in so called drill-and-kill-exercises about grammatical or lexical phenomena.

Some systems also provide a kind of guidance through the material of the system. In traditional programs there is a static system which is only able to recognize solved or unsolved exercises in order to provide guidance. This guidance is also part of the tutorial component of an intelligent program. The "tutor" suggests a certain way of working on the units and exercises, so that at the end the learner has covered all material following the given path. Usually this path is build following some pedagogical goals.

These two areas - exercise design and student modelling - are the main parts for improvement with the help of methods of CL.

- ***Exercises***: So far only pattern-matching with prestored examples is implemented in commercial programs and the content is usually static. Some kind of linguistic analysis can improve the feedback to the exercises. Nowadays this is possibly restricted to a orthographic or morpho-syntactic analysis. Further analyses in the semantic field would not yield good enough results for broad usage. The area of error analysis is further explored in the chapter "Error-Analysis", S. 43 .
- ***Guidance***: If some kind of learner-modelling is implemented the program can guide the learner according to his personal needs. Not only the performance in working on exercises can be measured but also the behaviour using the system like the frequency of using help-pages etc. In order to develop an intelligent tutoring system certain program-modules have to be realized as can be seen in the figure Grafik "ITS-Modules" . Most of these are necessary to provide adequate tutoring. More on this is explained in the chapters "Intelligent Tutoring Systems", S. 51 and "Learner Modelling", S. 56 .

  Based of this the help-system could also be improved. According to the performance of the learner, the system can provide tailored help-messages either in reference to the current exercise or to the handling of the program.

ICALL-systems which are tutoring systems include: PromisD Reuer, 2000, Promise Bauer,

John, Kronenberg, Krüger, Menzel, Reuer, Unsöld, 1994, CASTLE Murphy, McTear, 1997 (ReCALL-project[22]), Heift, 2001, CAVOL Kronenberg, Krüger, Ludewig, 1994 In all of these programs, the input of the learner is analysed with the help of CL technology.

22 http://www.infj.ulst.ac.uk/~recall/ - Last Visit: 6.2004

4.3. Functionality and Tools

Tools are used to find out something about a language. Therefore the relevance of properties of a language must also be taught. This coincides with new aims in language pedagogy. In these concepts learners are to discover the language by themselves. As mentioned before the learner follows an explorative approach to language learning. Programs used in this context are e.g. concordancers, which can help to find examples of language usage in "authentic" texts. This does not only support the learning of lexical aspects but also of grammatical and even of semantical aspects.

Note that the keyword "tool" does not refer to applications like a webbrowser or a wordprocessor, which can also be used in language learning. In order to improve "simple" applications such as concordancers the addition of CL-technology may be helpful. One example is the project TATOE[23], a concordancer able to use the output of the morphological analyser Morphy[24]. A morphological analyser can not only provide the input for further analysis tools but can be used directly in order to identify the morphological structure of words as it is done e.g. in Glosser[25]. Foreign electronic texts can then be read and understood more easily. The tagger might be able to classify the results or at least provide some more information about the structure of the found items. There are a number of tools for the analysis of spoken language possibly helping the language learner to improve his pronunciation. However I am not aware that this has been done in larger experiments. This might be a case where the learner needs an expert to explain to him the output of frequency-analysers and the like.

However one general problem is that learners generally do not need to learn abstract linguistic knowledge about a language in order to perform communicative actions. But as CL-tools usually deliver this kind of knowledge there are only very little areas where this makes sense.

Finally there is the area of style checking via assessment techniques such as "Latent Sematic Analysis" (LSA). In this case the tool may be used not for analysing language examples for the learner but to analyse learner language itself. There have been some experiments where texts written by learners are compared to a prototype text written by a teacher and thus evaluated and rated, see e.g. LSA at CU[26] (Summary Street project). The site also provides a number of introductory texts on LSA.

23 http://www.ipsi.fraunhofer.de/~rostek/tatoe - Last Visit: 6.2004

24 http://www.lezius.de/wolfgang/morphy/ - Last Visit: 6.2004

25 http://odur.let.rug.nl/~glosser/

26 http://lsa.colorado.edu - Last Visit: 6.2004

4.4. Functionality and Information Systems

As mentioned in the chapter on "", S. 27 the use of information systems is even more unidirectional than that of the previous examples. The only "interaction" is from the system to the learner. Therefore this subchapter means programs such as electronic dictionaries or grammars. The Chemnitz Internet Grammar[27] is a "simple" example with only little CL technology. However a corpus for exploration by the learner is included maybe counting as advanced compared to other online grammars.

Advantages of a WWW-based grammar are - on the one hand - access from everywhere and on the other hand the integration of advanced feature such as "free-text-search" and hyperlinks leading directly to related information. These cannot be found in printed textbooks.

A electronic dictionary has the same advantages as the grammar. Look-up can be done from anywhere and it can be faster than in a paper dictionary. Some features could include retrieval from inflected forms and the retrieval of corpus-items along with the basic dictionary-information. Additionally multimedia features can accompany the presentation of information.

Some word-processor have a thesaurus included, which can also function as an information system. A thesaurus usually only makes sense, if the learner has already some knowledge: When using alternative constructions one has to know the different connotations the various forms have. An example is the visualization of WordNet[28] at Visual Thesaurus[29]. Note that so far WordNet has only be created manually in order to exclude as many errors as possible. Glosser[30] includes a dictionary lookup over inflected forms with links to a corpus and a morphological analyser.

To sum up one can see many issues about functionality, which can be improved with the help of CL. All of this should of course be done with the learner in mind. Methods of CL can not only advance the possibilities of tutoring systems but also of tools and information systems to discover a language and thus teach and train a language.

27 http://www.tu-chemnitz.de/phil/InternetGrammar/

28 http://www.cogsci.princeton.edu/~wn - Last Visit: 6.2004

29 http://www.visualthesaurus.com - Last Visit: 6.2004

30 http://odur.let.rug.nl/~glosser/

4.5. Exercises Functionality

Homework in the summer course 2003

Name at least five possible methods of computational linguistics, and their application in CALL. Explain.

Clearcut questions

Name at least three areas, in which methods of CL can improve tutoring systems!

Outreaching questions

What would be the perfect feedback to get for a language learning exercise?

5. Error-Analysis

Abstract:

One of the main areas of applying CL methods to CALL has been the analysis of learner input. The idea is that the "deep" linguistic analysis of the input allows a precise characterization and adequate feedback of the errors the learner has possibly made. This has been tried on almost all linguistic levels from phonology to semantics. Additionally the deep analysis may allow an advanced type of exercise with almost free formed learner input. In this chapter the use of statistical methods such as the mentioned LSA for an evaluation of learner input is excluded.

5.1. Error-Analysis Phonology

Only very few advanced systems have been developed to recognize pronunciation errors with methods of NLP. In the ISLE-project[31] a system was developed, which is able to recognize mispronounced words and provide precise feedback about the error (some publications such as project-reports on the homepage[32]). Modern commercial systems nowadays also include speech recognition modules. However these systems do not analyse the input as such and can therefore only report about a percentage of correctness. Usually no hints about the type of error can be given, since the method solely relys on some form of pattern matching (comparing the frequency spectrum or the like). As a trivial example the program TriplePlay Plus which included a speech recognition module, presented a "ping" and "boing" as feedback when the pronunciation was close or could not be recognized. Other simple program show some sort of graph or frequency curve depicting the differences between the stored data and the learner recording. Note that the interpretation of the display is left to the learner.

The recognizer in the ISLE-prototype is a state-of-the-art HMM-based speech recognizer which is tuned towards the recognition of only a restricted set of words, but these are recognized with a high accuracy even if language learners pronounce the words.

In a first task the system tries to localize the error in the input. This is done in order to improve the error recognition task by sorting the input into correct and incorrect regions. The scores calculated depend on three type of measurements: 1. the acoustic likelihood of the path, 2. the output probility of the most likely state in the model set and 3. the acoustic likelihood of a background model.

The actual diagnosis follows in a second step. Some rules (letter-phone and phone-phone) are applied to the orthographic and the phonemic form of possible correct forms in order to determine the recognized form. A general description is as follows:

- articulatory difficulties producing particular sound of the target language (/th/ in English)
- receptive difficulties, because of which learners are unable to perceive and therefore to reliably produce the distinction between two sounds.
- orthographic carry-over from the mother tongue.
- orthographic difficulties of English.

Finally the system also includes a word stress detection mechanism. This allows the system to present feedback also in case the "simple" case of wrongly positioned stress was detected.

31 http://nats-www.informatik.uni-hamburg.de/~isle/

32 http://nats-www.informatik.uni-hamburg.de/~isle/

This description shows that on the one hand a very good speech recognition tool is necessary and on the other hand a rule-based mechanism is used in order to match the recognized string with the possible input.

5.2. Error-Analysis Morphology

Spelling correction is a standard feature nowadays in wordprocessors. Usually a word with the smallest so called editing distance is suggested for correction. Possibly for this reason there has not been much research in this area with regard to CALL. Note however that a simple spelling correction module is not useful for a learner of especially inflecting languages. Usually there are some default inflection-paradigms, which are learned by the student and are then applied according to his knowledge. The problematic cases are the nondefault ones, which an intelligent system should correct accordingly. A program should ideally not only state that "er gehte" is wrong and may be corrected to "er geht" (dropping the final "e" as the smallest change to the original), but it might present a message saying that "gehte" is wrong because the verb "gehen" is irregular and the correct past tense form is therefore "ging".

What is required in these cases is a module using the morphology of a language to determine the cause of errors from inflection paradigms.

There has been some research on the relation between the phonology and the orthography of a language as it is usually assumed e.g. for German, that there is a certain close relationship between these two areas. A common approach is the Soundex approach (Mitton, 1996). The idea is that a misspelled word not in the lexicon is reduced to a highly simplyfied phonological form via a small set of rules. A comparison with the reduced representation of correct words should reveal a set of candidates. This of course only works if the string in question is definetly not in the lexicon and the spelling error is actually based on a phonologically similar form.

Finally there has been a concept developed for a mapping of the German syllabel system and orthography (Maas, 1992). Based on this a program was developed not for foreign language learners but for young children learning to write. The ruled based system produces exercises training the recognition of the internal structure of words and the corresponding spelling especially for lenghend and shortend syllabels.

5.3. Error-Analysis Syntax

Two general strategies for morhosyntactic error handling can be distinguished. Firstly there are so called *robust* parsing methods, which try to continue parsing past a position, which cannot be handled by the grammar without considering the type and exact location of the error. The main purpose is to achieve a result for as much of the input as possible, which usually means to yield the largest possible chunks. A similar approach is used for analyzing spoken language, where additionally the so called recognizer may detect erroneous structures such as interruption or repetition of a phrase, corrections etc. Secondly *sensitive* strategies are being developed, which specifically try to locate and analyze errors in the input. With the help of some type of correction method the parsing process will continue across the error position and yield a complete description of the input usually including the position and the type of error. In a system which aims at determining the grammaticality according to a given grammar and at providing as much feedback about an error as possible only the second type of parsing method can be adopted.

Again, two strategies can be followed for identifying errors: Either the algorithm is changed to allow for the recognition even though the grammar does not cover the input, or the grammar is extended with so called mal-rules, which allow the generation of a description. The first concept can be refered to as "anticipation less" whereas the second one is called "anticipation-based". The following discusses a few aspects of these two approaches to error recognition. An example from Schwind, 1994 can demonstrate one major disadvantage of the second approach. The mal-rule in the phrase structure grammar is specifically designed to describe an error a French native speaker would make when learning German, namely to position the adjective after the noun in a noun phrase (le maillot jaune vs. das gelbe Trikot). However speakers with other mother tongues may produce other errors not covered by this approach. New rules would have to be added to the grammar for many other cases. A similar case is presented in Schneider, McCoy, 1998, where a grammar rule is designed to specifically recognize a number mismatch between the determiner and a noun in a noun phrase. Referring to agreement values in PS-rules increases the problems even more, because covering a substantial grammar fragment would cause the number of rules to explode. This in turn leads to enormous efficiency problems.

However two advantages have to be noted with regard to the anticipation-based approach. One is, that the most efficient parsing algorithms can be chosen, since usually the grammars are not changed in their form but only extended. A second advantage is the possibility to be able to distinguish between on the one hand ungrammatical input and on the other hand unparsable input, i.e. input, which is not covered by the grammar. In most cases this means that the feedback to the learner can be stated with more confidence about the location and the type of the error.

The anticipation-free approach transfers the load of recognizing and handling errors in the input into the parsing mechanism. An example of this is the approach presented in Menzel, 1992, where a "model-based" error-diagnosis is characterized by its complexity. For example in an agreement situation every feature of every lexical item is checked by an individual function in order to allow for a precise localization of an error. As another example for the load moved into the parsing mechanism the approach taken by Mellish, 1989 needs two parses in order to identify linearization errors with the help of a simple phrase structure grammar. One important advantage is the recognition of certain errors "anywhere" if they can be identified at a single position. As an example an agreement error between the subject and a verb should be recognized not only if the subject is in the standard position but also if for some reason the subject is displaced, e.g. by topicalization. A second advantage is the chance for independent development of a grammar and a lexicon. They can be engineered so as to generate descriptions only for correct sentences of a language. This also allows the integration of "foreign" data, e.g. a large lexical database, which may have been developed in other contexts.

In order to decrease the processing load, the search space for finding a solution may be minimized in anticipation-free concepts. The evaluation in Lee, Kweon, Seo, Kim, 1995 can be taken as an example for this problem (for a certain type of robust parsing): With a phrase structure grammar with only 192 rules, the parsing algorithm generates 12.000 items in the chart with heuristics turned on and even 25.000 items without. On the one hand almost any sentence will be analyzed but on the other hand the efficiency is very low. In order to counter this general problem Mellish introduces a number of heuristics, which refer to a variety of possible configurations of chart items. One of the main aims of Kato, 1994 with his modification of the algorithm is therefore to decrease the number of different heuristics and nevertheless improve the efficiency. A different approach is chosen in Schröder, Menzel, Foth, Schulz, 2000 and Fouvry, 2003. In these two cases constraints are weighted, which on the one hand allows robust parsing and on the other hand nevertheless allows determining the solution with the "smallest" error measure. Additionally in Schröder, Menzel, Foth, Schulz, 2000 constraints can be marked with a weight 0, which effectively makes them so called hard constraints. Solutions with this kind of constraint clash will then not be considered for the further analysis. However there is no evaluation with respect to a possible feedback to the learner about the error in these two approaches.

5.4. Error-Analysis Semantics

In order to determine an error a system needs some kind of description of what is correct. For semantic errors this amounts to a (almost) complete description of the possible states of the world including all things in the world. This task has not been realized as yet which makes it difficult to identify semtic errors in general. However limiting the domain in a sensible way can provide a system with some capabilities for recognizing semantic errors. The presentation of a so called micro world is a step towards this goal.

In the SEMF-project[33] the system tried to recognize not only syntactic errors but also semantic errors. Some restrictions need to be added to realize such an ambitious task. In the example here the input sentences are always decriptions of a so called blocks world. After parsing a sentence syntactically the system then checks if the semantic representation of the sentence matches the actual blocks world at that moment. Since the contents can only refer to the items in the blocks world this seems to work. An application of this method to larger and more complicated worlds seems albeit difficult. A complete representation of a micro world would have to be developed not only with states but possibly also including changes of states (but see Hamburger, 1994, Hamburger, 1995, who proposes a kitchen scene manipulated by the learner).

Some more basic semantic knowledge such as thematic restrictions may already be covered in the system's lexicon. This approach has been tried in the ReCALL project. The grammar rules were modified in order to include semantic information parallel to the usual syntactic agreement information. This could then be used in order to generate appropriate feedback to a learner if the restrictions were violated. Again these constraints can only be used in very restricted cases and are not a general method of identifying semantic errors.

To summarize almost all linguistic levels have been considered in research on error-analysis. Most has been done on grammar checking, which on the one hand seems to be managable and on the other hand is important for learners. Little work has been done on the presentation of error-messages to learners (but see Heift, 2001). This second important step has not been considered much maybe because analysing errors is already a very difficult task. As mentioned above, there is some research from the teaching perspective, e.g. Gnutzmann, Kiffe, 1993. Here the question is, which type of error-handling students prefer while speaking in foreign-language classes.

33 http://www.cogsci.uni-osnabrueck.de/~semf/

5.5. Exercises Error Analysis

Homework in the summer course 2003

Read two papers: Schneider,McCoy 1998 (Coling) and Menzel 2002 (Kognitionswissenschaft; in German) about Error Recognition. Both are available online and in the ILIAS-database. Write down 5 questions you have about each text and send them in before the next session.

Freshen up your knowledge about Computational Semantics.

Read the paper Menzel 2001 (ReCALL Journal) available online.

6. Intelligent Tutoring and User Modelling

Abstract:

In order to provide an individualized learning environment the CALL systems needs to be able to adapt the system to the learner. This is done on some general assumptions about language learners in general and also on specific information collected during the interaction of the learner with the system. A number of program modules are suggested to process the information from the learner, to draw conclusions from it and to adapt the system accordingly. A central aspect is the learner's model and the possibilities of adaptation.

6.1. Intelligent Tutoring Systems

Some parts of this section are based on Schulmeister, 1997 and Beck, Stern, Haugsjaa, 1996. An introduction to ITS can also be found in e.g. chapter 6 of Schulmeister's book. With the development of Artificial Intelligence in the 60s there was an immediate adaptation of AI-techniques for ITSs. One of the first systems considered an ITS was developed by J.R. Carbonell and was called SCHOLAR Carbonell, 1970. Its aim was to teach South American geography. The system had a tutoring-module which was able to infere the appropriate steps through the program from the student's responses. One important development in this case was the clear division between the inference engine and the domain knowledge. These two modules are already some of the necessary components of an ITS. Usually one considers the following parts:

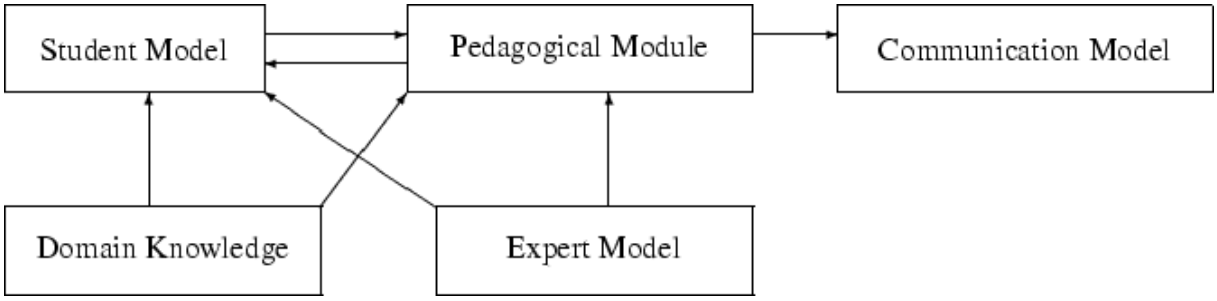

Figure: ITS-Modules

Domain Knowledge

The domain knowledge usually consists of declarative, procedural and sometimes heuristic knowledge:

- Declarative knowledge defines terms with their attributes and defines the relationship of the terms to each other.
- Procedural knowledge consists usually of rules (i.e. *how* to do things), which aid the problem-solving-process.
- Heuristic knowledge can be seen as the experience an expert would use for solving problems. Heuristics can be used for faster and sorted analyses.

Basically two models for the domain knowledge have been assumed in research. There is a black-box-model on the one hand which makes no claim about human behaviour regarding inference. The glass-box-model on the other hand claims that it reproduces intelligent reasoning and structuring of information. More about this topic can be found in e.g. Goerz, Rollinger, Schneeberger (Hrsg.), 2000.

In an ICALL-system the knowledge would cover mainly the linguistic knowledge necessary to be able to use a language. This might include a grammar and a lexicon, but also a corpus might be added to this list.

Expert Model

The expert model is similar to the domain knowledge. It is different in that it is a model of how an expert would use the knowledge contained in the domain knowledge. Thus it is more than just the representation of knowledge. For the assessment of a solution from the learner the solution can be compared to the solution the expert model provides on the same domain knowledge. Some models of ITS leave out this part and consider it included in the domain knowledge.

Student Model (or learner model)

The learner model is build in order to allow the system to refer to the probable current knowledge of the learner. The pedagogical module mainly uses this information. The system tries to determine the current knowledge either on the basis of a *subset* of the domain knowledge or of the *deviation* of the learner-performance from the expert performance in the same situation (overlay model). Sometimes Bayesian networks are used to determine the student's knowledge probabilistically, where each node in the network has a probability to indicate the likelihood of the student knowing this. In the learning-process any input from the learner is analysed and from this the knowledge of the learner is concluded. This does not take into account the general learning behaviour, as the actions from the student can only be observerd in "a narrow channel". Sometimes it will be difficult to assess whether the learner does not know something or whether a different strategy than the expert model suggests was used to acquire the knowledge.

The knowledge contained in the learner model can theoretically vary between the very general and simple statement "the learner knows this domain" and a detailed listing of all the actions the learner has taken. Of course most systems lie inbetween, i.e. some general and some specific information is used to represent the learner's knowledge. They mostly use the same granularity which is used in the domain knwoledge base.

Self, 1988 determines 6 groups of functions a learner model can have.

- Corrective function: The system must be able find an error and to correct the learner if he has made an error.
- Elaborating function: If the system decides, that the learner's knowledge is correct but incomplete it must point out this fact to the learner and suggest actions.
- Strategic funtion: The analysis of the learner's performance might lead to different learning strategies.
- Diagnostic function: The system must be able to decide on certain actions by itself.
- Predictive function: Simulation of the learner's behaviour should lead to predictions about further actions.

- Evaluative function: The system must be able to backtrack and reconstruct the "learner's learning process" and evaluate the learner's performance on the basis of this reconstruction.

For diagnosing errors most system rely either on error-libraries or some method of machine-learning to be able to identify errors. A special problem is that errors from learners tend to occur not as single errors but compounds adding difficulties to the analysis process. Usually some heuristics are developed to minimize the search-space of hypotheses for the errors.

Schulmeister, 1997 explicitly mentions the missing psychological principles used to evaluate the psychological plausibility of a solution. This is something which he sees as an important step towards reliable performance-evaluation.

Pedagogical Module (or tutor module)

The tutor model is responsible for the presentation of the right learning material at the right time. It therefore simulates the teacher in deciding on the basis of the expert knowledge, which action to take next. Since the student model is also used as input to this module the different actions reflect the different needs of each student.

Two main tutoring strategies are used in most systems.

- The Socratic dialog: The tutor asks questions and tries to analyse the learner's input to build up a learner model. This means the system is responsible for planning the steps through the program.
- The coaching approach: The learner is allowed to exercise his/her skills and try out different solution. The system only reacts to input "when it is asked to do so". Note that this approach is much closer to the already mentioned didactic principle of explorative learning.

After the decision on a "meta-strategy", low-level aspects must be considered. Following Beck, Stern, Haugsjaa, 1996 this includes:

- Topic selection: This includes the selection of new material or of old material for reviewing.
- Problem generation: Following the selection of a topic a problem for the learner to be solved must be generated or selected. This is in great parts based on the student model.
- Feedback: Especially when the learner makes an error, the type and style of feedback is important. The advice given to the learner should be "appropriate to the ability level". Ideally the system chooses feedback as much as necessary and as little as possible.

What is missing in this respect in a lot of systems is what Schulmeister calls the "grammar of interaction". The system usually does not have knowledge about the structure of the situation

and of general rules of interaction. Finally criticism is brought forward because "students are never involved in the development of tutor models", which leads to rather abstract models of students.

In this module the various task-types and their content are organized. For example in an ICALL system this module should decide whether a vocabulary test or a grammar exercise at which level is to be presented to the learner.

Communication Model

The communication model is the model that appears as the "intelligent" model to the user. The system reacts flexible in the presentation of learning materials and adaptive with regard to the assumed knowledge of the learner. Different forms of interaction are possible:

- The system puts questions to the learner, guides him by analysing the answers and posting new questions.
- The system lets the learner work for himself and waits until a question is presented.
- The system is active itself and "invites the learner to select information" and infers from the selection the assumed knowledge.
- The system remains in the background and only offers help now and then.

In general there are usually two types of systems. There is the system which follows the instruction-concept. This is the more guiding system which does not allow the learner to move around freely in the learning environment. It has a lot of interaction with the learner in order to guide the learner through units and exercises. The microworld-concept (or construction-concept) on the other hand is more free and does not much interfere with the learner's actions. This calls for making hypotheses by the learner about the material and experimenting with these hypotheses in the learning process.

It is not clear whether a natural-language-interface is really necessary to make up an ITS. Some authors claim that an ITS should at least realize a dialog which comes as close as possible to a natural dialog. This point must not be confused with language learning. Language is simply the main instrument for communication and therefore should be used as much as possible to allow the learner to express himself in the way he is used to.

The main reason for choosing an ITS is still the argument of individualization. An ITS should be able to adapt itself so that the learner is ideally supported for his learning needs. "Learning errors" like flaundering or overlooking learning opportunities are at the aim of an ITS to be avoided.

A major difficulty of ITSs is the gap between cognitive concepts which can be modelled according to the needs of the computer and pedagogical and psychological theories which have a different methodological status according to Schulmeister, 1997. Hence they can not be

mplemented in a computerprogram. One major problem is the meaning of "understanding". Usual definitions of "understanding" can not be applied to ITS. True cognitive systems which might be able to "understand", have not been developed yet. The difference can also be described as "cognition" being on the one hand in cognitive psychology and on the other hand in cognitive science.

The use of simpler models of learning in ITS than the ones in cognitive psychology is simply the hope for theories, which can one day be implemented in a computer.

Adaptivity - Dimensions for categorization

Some parts of this section are based on Beck, Stern, Haugsjaa, 1996. One dimension can be the level of simulation a program tries to achieve. Some systems are simulations, which try to cover a most realistic working part of the real world whereas other programs might be further away from reality. Some programs even teach in a "decontextualized" manner. This then constitutes the opposite extreme in this dimension.

Another dimension is the type of knowledge, which is being taught. There have been several attemps of classifying "educational objectives" cf Bloom (Hrsg.), 1956. One could also ask what the student will be able to do when he has learned all the material from the Unit. This is a similar question but is not equal to the previous one. On the one hand the learner might be able to perform skills and on the other hand an abstract theory is now known to the learner. The most common type of knowledge is some kind of procedural skill. The learner should be able to perform a particular task. Since some systems are based on research in the cognitive psychology of human skill acquisition they may be called *cognitive tutors* Beck, Stern, Haugsjaa, 1996. This is usually the case, when in some way the analysis of the learner action is based on a "cognitively" plausible expert model.

Knowledge based tutors in contrast to the former type are systems which aim to teach things like concepts and "mental models" (frameworks). Since there is less knowledge of how concepts are acquired by humans these systems tend to have a lot more domain knowledge. General teaching strategies are used to present the right material to the student at the right time. With the help of the knowledge base explanations to the student's action are generated. In both types of dimension the actual systems must be seen along a continuum. There are e.g. sytems which teach how to use Email-programs in UNIX, which are procedural skills. But then a large knowledge base is used to develop the explanations and the feedback.

There have been a series of conferences called ITS 'Year, for which proceedings were printed. The last one was 2000 in Montreal in Canada.

Books, articles etc.: Sleeman, Brown (Hrsg.), 1982, Woolf, 1987, Sama Nwana (Hrsg.), 1993, Yazdani (Hrsg.), 1993, Greer (Hrsg.), 1994

6.2. Learner Modelling

This chapter contains information about user modelling in CALL systems as opposed to the previous more general chapter. One aspect in this area is the need for collecting information about the user, i.e. the language learner. Usually this can be done by "watching" the user interact with the system selecting pieces of information. In the language learning scenario there is the additional chance to collect information from the learner about his performance while working on exercises. As mentioned in the previous chapter a learner profile can then be generated. Therefore there are two main areas for *using* a learner profile in CALL systems. One is the adaptation of the system itself towards the user and the other one is the generation of "fitting" exercises.

One project, which aims at doing learner-modelling in a dictionary-setting is ELDIT[34]. In the system 5 features can be adapted according to the learners needs.

Feature	Setttings
model	monolingual/semibilingual
domain	general/medical/technical
difficulty	beginner/advanced
help	novice/familiar
pronunciation	local/standard

Customizable features in ELDIT (taken from Gamper, Knapp, 2001)

From table "", S. 56 one can see, that there are various possibilities to adapt the program to the user's needs. Two main areas can be identified.

1. The general aspects of HCI. A system can be adapted to needs and to the preferences of a learner in the simple tasks of interaction. This can be e.g. the language used for buttons, hints and help-texts. Some learners might prefer their L1 whereas more advanced learners might prefer the foreign language.

2. The presentation of learning materials. An ICALL system should ideally adapt to the assumed knowledge of the learner about the foreign language he is learning. This can only be done of course, if either the learner places himself on a proficiency-scale or the system analyses some exercises solved by the learner.

Another project is the ICICLE[35]-project. This project aims at deaf native speakers of American Sign Language (ASL) to help them learn writing. In the program user-modelling is done according to possible language acquisition models. Comparing the learner performance of the learner with the acquisition model the system can e.g. establish hypothesis about the know-

34 http://www.eurac.edu/eldit

35 http://www.eecis.udel.edu/research/icicle/

ledge of the leaner.

A second aspect of learner modelling in this project is the more precise evaluation of language input. Using the system the learner enters a sentence, which is parsed, i.e. morpho-syntactically analyzed. But because more than one parse may result, a learner model helps to select the most viable one. In the ICICLE-project this is done according to the so called overlay-model.

Every grammar rule in the grammar the parser uses is marked according to a certain proficiency level the learner can reach. With the help of these markings and a seperate expert model the system can provide some information about the proficiency of the learner. This also helps in determining a parse result to choose. Furthermore the system updates the information about a certain learner.

6.3. Exercises Intelligent Tutoring and User Modelling

Homework in the summer course 2003

What kind of dimensions can be used in learner modeling in ICALL-systems, e.g. beginner/ advanced? In which ways can a program be adapted to a learner?

Special presentation: Develop an introductory bibliography on ITS with a focus on language learning. Include resources from the WWW.

Read two texts: Gamper,Knapp 2001 (ABIS) and Michaud, McCoy, Stark 2001 (UM). Think of 5 questions you have about the texts each.

Mixed Questions

State in your own words the difference between procedural and declarative knowledge?

To which category does language belong?

Can you think of other models and modules an ITS would need?

If a learner avoids the usage of something, does he not know it?

7. Resources

Organisations

EUROCALL[36] and the

SIG Language Processing[37]: Contains a large bibliography about CALL including ICALL.

SIG Integrating Speech technology in Language Learning[38]

CALICO[39]

IALLT[40]

AIED[41]

Journals

Computer Assisted Language Learning[42]

Language Learning and Technology[43]

CALL-EJ Online[44]

ReCALL[45]

Project Homepages

36 http://www.eurocall.org

37 http://siglp.eurocall.org

38 http://dbs.tay.ac.uk/instil/

39 http://www.calico.org

40 http://www.iallt.org

41 http://www.cbl.leeds.ac.uk/ijaied/aiedsoc.html

42 http://www.swets.nl/sps/journals/call.html

43 http://llt.msu.edu/

44 http://www.lerc.ritsumei.ac.jp/callej/

45 http://www.eurocall.org/recall.htm

FreeText[46]

ICICLE[47]

ISLE[48]

LogoTax[49]

ReCALL[50]

Other Weblinks

ICT4LT[51]: `This website on Information and Communications Technology for Language Teachers is the result of over two years' work by an international team of experts during the period September 1998 to December 2000.' Contains also a large bibliography and further links.

Some texts[52]

Some programs[53]

Demo-programs from Heureka-Klett[54] (> Produkte > Softwareproben)

Catalogue from Hueber[55] (> Service > Katalog)

Catalogue from Cornelsen[56] (> Eltern: Katalog plus > Lernsoftware (Cookies!))

46 http://www.latl.unige.ch/~freetext/

47 http://www.eecis.udel.edu/research/icicle/

48 http://nats-www.informatik.uni-hamburg.de/~isle/

49 http://cato.cl-ki.uni-osnabrueck.de/~logotax/

50 http://www.infj.ulst.ac.uk/~recall/

51 http://www.ict4lt.org

52 texts/

53 programs/

54 http://www.klett-verlag.de/heureka/

55 http://www.hueber.de

8. WWW-accessible programs

Abstract:

In this chapter three example systems are listed. On the one hand the all of them are accessible via the WWW and on the other hand they represent the different categories used in the previous chapters. Ther parser from the PromisD project is used in a tutorial system. The LogoTax system can be considered a tool as it serves as a personal electronic lexicon and finally Glosser is an information system providing data on a fixed set of texts.

8.1. PromisD

ErrLFG[57]: the parser from PromisD

Some questions

What is an ungrammatical utterance?

How can syntactic errors be categorized?

What types of errors can ErrLFG detect?

How is an error encoded in the LFG-structures?

[57] http://cato.cl-ki.uni-osnabrueck.de:2212/~vreuer/lfg/

8.2. LogoTax

LogoTax[58]

Some questions

What are collocations?

Why are collocations important in foreign language education?

How are collocations treated in traditional dictionaries?

What are the differences between Online-dictionaries such as Pons and LEO and paper-dictionaries?

How are collocations described in LogoTax?

Which description levels are used?

[58] http://cato.cl-ki.uni-osnabrueck.de/~logotax

8.3. Glosser

Glosser[59]

Some questions

What kind of lexical information can be provided automatically with the help of CL?

What kind of information does Glosser provide?

Why is the program called "Glosser"?

[59] http://odur.let.rug.nl/~glosser/

Bibliography

Articles

Ahmed, Mehreen: *A Note on Phrase Structure Analysis and Design Implication for ICALL*. In: *Computer Assisted Language Learning*. 15 (4), 2002, pp 423-433
Exploiting Chomsky's Universal Grammar this research examines the effect of phrase structure rules on the development of grammatical knowledge in Bangladeshi learners of English as a second language. The masking technique of McWhinney motivates the use of the link-grammar syntactic parser of Sleator and Temperley used in the study. The results indicate that the learners' parameter were reset from L1 to L2 as a consequence of the overt presentation of phrase structure rules. The implication of the study for the design of intelligent computer-assisted language learning systems is discussed. The results however are drawn on rather weak evidence as only observations, interviews and open-ended questionaires were used.

Bayerlein, Oliver: *Versuch einer Übungstypologie für computergestützte Multimedia-Sprachkurse*. In: *Info DaF*. 23 (6), 1996, pp 726-736
This article about exercise-types and the possibilities of categorization show an interesting approach. The main categories are: Verbale Semantisierung and Integration und Automatisierung. One category is explained through the function, the media and an example. Altogether 12 categories are distinguished.

Beck, Joseph ; Stern, Mia ; Haugsjaa, Erik: *Applications of AI in Education*. In: *ACM Crossroads*. 3 (1), 1996
http://www.acm.org/crossroads/xrds3-1/aied.html - Last Visit: 5.2004
The paper provides an overview of the main components af intelligent tutoring systems and a brief summary of different types of ITSs. Next a detailed discussion of two components, the student model and the pedagogical module is presented. Finally a few open questions in ITS are discussed. Even though this text is a few years old it provides a comprehensive although short overview of the main modules of a generic ITS.

Breindl, Eva: *Grammatik im WWW*. In: *ORBIS: Online Reihe Beiträge zu Sprache und Internet*, 1999
http://www.ids-mannheim.de/grammis/orbis/tswww/tsframe1.html - Last Visit: 7.2000

Carbonell, Jaime R.: *AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction*. In: *IEEE Transactions on Man-Machine Systems*. 11 (4), 1970, pp 190-202
This article is one of the very early texts about the application of methods of Computational Linguistics and Artificial Intelligence to CALL. It discusses the AI techniques from that time

which make it possible to incorporate meaningful tutorial dialogues in computer assisted instruction programs. It also describes the SCHOLAR program, which incorporates an information network of facts, concepts and procedures, which enables it to carry on a mixed initiative dialogue between student and computer, handling questions and answers from both sides. The content is geographical issues in South America.

Eidecker, Martina Elisabeth: *Anwendungsweisen multimedialer Computerprogramme im Fremdsprachennterrricht*. In: *Deutsch als Fremdsprache*. 2, 1996, pp 109–115
CyberBuch ist eine multimediale Software, die Studenten den Zugang zu authentischen Lesetexten erleichtern soll. Sie unterstützt die Entwicklung von Lesestrategien, indem sie nicht nur den Text anbietet, sondern auch „Organisatoren", die den Text kontextualisieren (Filme, Fotografien und Hintergrundinformationen). In Übereinstimmung mit dem Natürlichen Ansatz in der Spracherwerbstheorie finden Sprachaneignung und Spracherlernen nur in der Zielsprache statt. Über die Vorstellung dieses Programmpakets hinaus wird aber auch kaum ein neuer Aspekt von Computer und FU eingebracht. Aus heutiger Sicht ist der Artikel damit zum großen Teil veraltet, da es nun viele CyberBuch ähnliche System gibt.

Gamper, Johann ; Knapp, Judith: *A Review of Intelligent CALL Systems*. In: *Computer Assisted Language Learning*. 15 (4), 2002, pp 329-342
This paper provides an overview about intelligent computer-assisted language learning. The most advanced systems have been investigated and classified along five dimensions: supported languages, AI techniques, language skills, language elements and availability. The paper concludes with a discussion about outstanding problems. All the important systems after 1995 are included and a good overview is given in a comprehensive table. Unfortunately only one major publication per project is mentioned which limits the the bibliography a bit.

Gnutzmann, C. ; Kiffe, M.: *Mündliche Fehler und Fehlerkorrekturen im Hochschulbereich. Zur Einstellung von Studierenden der Anglistik*. In: *FLuL*. 22, 1993, pp 91-128

Goettmann, Hans: *Schreiben und Üben mit dem PC*. In: *Info DaF*. 23, 1996, pp 69-80

Granger, Sylviane ; Vandeventer, Anne ; Hamel, Marie-Josee: *Analyse de corpus d'apprenants pour l'ELAO base sur le TAL*. In: *T.A.L.*. 42(2), 2001, pp 609-621
In this article Granger et al. report about the FRIDA corpus of learner's French and the combination with the FreeText project in which a grammar checker was developed. First the construction and annotation of the corpus is described and in the second part an evaluation of the parser used in the FreeText project is presented. Finally some conclusions with regard to the usage in language learning and didactic issues are discussed. Since a lot of topics need to covered there are not too many details in the text. But it gives a good impression of the chances of using certain types of CL technology for CALL.

Healey, Deborah ; Johnson, Norm: *A Place to Start in Selecting Software*. In: *CAELL*. 8(1), 1997, pp 3-9
http://oregonstate.edu/~healeyd/cj_software_selection.html - Last Visit: 5.2004
This text introduces a few aspects to consider when planning to use a CALL-program. There is a set of six questions with explanations to ask oneself before choosing a software. In the second part the table with software recommendations is explained.

Heift, Trude: *Error-Specific and Individualized Feedback in a Web-based Language Tutoring System: Do They Read It?*. In: *ReCALL*. 13 (2), 2001, pp 129-142
In this article Heift presents a user study with 33 students using her system E-Tutor. Even though the exercises are fairly restricted the systems uses a HPSG-based parser for the generation of "meta-linguistic" feedback. The learners were able to decide whether they wanted to read specific feedback or leave it at general remarks. The study shows that the student actually read the more specific texts and even paid increasing attention to the feedback with more iterations of the exercises. Even though no "post-test" was conducted the study shows that students seem to pay attention to the parser generated feedback.

Heift, Trude ; Nicholson, D.: *Web delivery of adaptive and interactive language tutoring*. In: *International Journal of Artificial Intelligence in Education*. 12(4), 2001, pp 310-325
This text contains the main description of the E-Tutor system developed by Trude Heift et al. Notably it describes the step wise analysis of an input sentence in order to achieve a robust and meaningful analysis. Also an adaptive student module is described which especially adapts the contents of the error messages to the learner's needs.

Hess, Hans W.: *DaF-Software in der Anwendung - Alter Quark noch breiter?*. In: *Info DaF*. 25, 1998

Jensen, K. ; Heidorn, G.E. ; Miller, L.A. ; Ravin, Y.: *Parse Fitting and Prose Fixing: Getting Hold on Ill-formdness*. In: *Computational Linguistics*. 9(3-4), 1983, pp 147-160
http://acl.ldc.upenn.edu/J/J83/J83-3002.pdf - Last Visit: 6.2004
This paper describes the EPISTLE system and the error detection capabilities. For syntax errors there is a two stage process, in which first a sentence is parsed bottom up until no further analysis can be produced any more. In the second phase a set of so called fitting rules is applied in order to achieve some interpretation of the sentence. Therefore a sentence is not really corrected but the parser is able to produce some kind of analysis for further processing. However the system also includes correction procedures decribed elsewhere. Finally the system also includes a style checker which uses some heuristics such as repetition in order to suggest corrections.

L'haire, Sebastien ; Vandeventer Faltin, Anne: *Error Diagnosis in the FreeText Project*. In: *CALICO Journal*. 20(3), 2003, pp 481-495
In this article the authors present the results from the FreeText project. The system developed has a different approach of error diagnosis on every linguistic level which are presented level by level (orthography, syntax and semantics). Based on an already existing large coverage parser the syntactic module is able to identify quite a number of errors. Note also, that here one of the few approaches to tackling semantic errors is presented.

Ludewig, Petra: *LogoTax - un outil exploratoire pour l'etude de collocations en corpus*. In: *T.A.L.*. 42(2), 2001, pp 623-642
In this paper the LogoTax system is presented. LogoTax is a kind of personal electronic dictionary which supports the learner in entering items into a personal database. To reach this goal corpora are analysed syntactically in order to provide feedback especially for collocations. The learner should be able to identify the usage patterns of a collocation through the corpus examples which additionally are presented after a syntactical evaluation. This process is described indepth. Note however that the system can also store any other kind of entry the user whishes to keep.

Lyon, Gordon: *Syntax-Directed Least-Error Analysis for Context-Free Languages: A Practical Approach*. In: *Communications of the ACM*. 17(1), 1974, pp 3-14
This paper describes a parsing approach based on Earley-style chart parsing which can correct various types of constituent errors. Planned for the parsing of programming languages it was taken up by Lee et al. (1995) for parsing natural language. Errors are detected by adding chart items representing error hypotheses on a local level. This increases the search space but allows to identify quite a number of errors.

Menzel, Wolfgang ; Herron, Daniel ; Morton, Rachel ; Pezzotta, Dario ; Bonaventura, Patrizia ; Howarth, Peter: *Interactive Pronunciation Training*. In: *ReCall*. 13(1), 2001, pp 67-78
In this paper methods and modules are described for the automatic localisation and correction of pronunciation errors developed in the ISLE-project. This work was part of the project aimed at integrating state-of-the-art speech recognition technology into an pronunciation training environment for adult, intermediate level learners. The system focuses on Italian and German learners of English although in principle any language pair is possible. The article actually gives an overview and therefore does not describe the components with too much detail as the focus is especially on the feedback.

Nagata, Noriko: *Intelligent Computer Feedback for Second Language Instruction*. In: *The Modern Language Journal*. 77(3), 1993, pp 330-339
In this article one of the few projects is presents that actually tried to measure whether students learned better when provided with detailed analyses from an ICALL system. The

experiment was done with English learners of Japanese for certain constructions of Japanese. The result show a marginal improvement for the learners with the advanced feedback. Note again that no comparable study exists to my knowledge.

Oflazer, Kemal: *Error-tolerant Finite-state Recognition with Applications to Morphological Analysis and Spelling Correction*. In: *Computational Linguistics*. 22(1), 1996, pp 73-89
http://acl.ldc.upenn.edu/J/J96/J96-1003.pdf - Last Visit: 6.2004
This paper presents the notion of error-tolerant recognition with finite-state recognizers along with results from some applications. The only precondition is the availability of a single and possibly very large finite-state transducer. Misspelled words are corrected if there are candidates with a certain edit distance from the misspelled word.

Reuer, Veit: *Error Recognition and Feedback with Lexical Functional Grammar*. In: *CALICO Journal*. 20(3), 2003, pp 497-512
This paper describes the error recognition module of an interactive ICALL system with a special focus on the underlying grammar theory. It is argued 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. The concepts and structures used in LFG closely resemble the descriptive knowledge of language learners about a language, and, therefore, the results of an automatic analysis can easily be translated from a computationally tractable form to language easily understood by the learner. Note however that the formal aspects of the parsing algorithm are left out sometimes in order to highlight the possibilities for the actual learner.

Rösler, Dietmar: *Autonomes Lernen? - Neue Medien und altes Fremdsprachenlernen*. In: *Info DaF*. 25, 1998

Salaberry, M. Rafael: *Review of Sake Jager, John A. Nerbonne, A. J. van Essen (Eds.): Language Teaching and Language Technology*. In: *Language Learning & Technology*. 4(1), 2000, pp 22-25
http://llt.msu.edu

Schröder, Ingo ; Menzel, Wolfgang ; Foth, Kilian ; Schulz, Michael: *Modeling Dependency Grammar with Restricted Constraints*. In: *T.A.L.*. 41(1), 2000, pp 113-142
In this paper, parsing with a dependency grammar is modeled as a constraint satisfaction problem. A restricted kind of constraints is proposed, which is simple enough to be implemented efficiently, but which is rich enough to express a variety of grammatical well-formedness conditions although with some limitations. A number of examples is given to demonstrate how different kinds of linguistic knowledge can be encoded in this formalism.

Schulze, Mathias ; Hamel, Marie-Josee ; Thompson, June: *Language processing in CALL*. In: *ReCALL*, 1999
Special Issue

Schwind, Camilla: *Error Analysis and Explanation in Knowledge Based Language Tutoring*. In: *Computer Assisted Language Learning*. 8(4), 1995

Segler, Thomas ; Pain, Helen ; Sorace, Antonella: *Second Language Vocabulary Acquisition and Learning Strategies in ICALL Environments*. In: *Computer Assisted Language Learning*. 15 (4), 2002, pp 409-422
The research described in this paper investigates the role of Vocabulary Learning Strategies (VLS) in ICALL environments. The paper does not present finished results but rather tries to assess which strategies to consider and how vocabulary acquisition could be evaluated. Therefore the paper is rather an overview on strategies for a certain concept or project.

Vandeventer, Anne: *Creating a grammar checker for CALL by constraint relaxation: a feasibility study*. In: *ReCALL*. 13(1), 2001, pp 110-120
This study in the context of the FreeText project explores the usability of the Fips parser for error recognition. The paper reports on the recognition of agreement errors in French sentences taken from the FRIDA corpus also mentioned in Granger et al. 2001. The specific technique used is a kind of constraint relaxation. The relevant feature receives a marking that an error has been encountered. In total a precision above 50 per cent and a recall above 70 per cent was reached. However a large number of false positives were encountered which is discussed in detail. Note that the technical details are only sketched.

Weischedel, Ralph M. ; Voge, Wilfried M. ; James, Mark: *An artificial intelligence approach to language instruction*. In: *Artificial Intelligence*. 10, 1978

Weischedel, Ralph M. ; Sondheimer, Norman K.: *Meta-rules as a Basis for Processing Ill-Formed Input*. In: *Computational Linguistics*. 9, 1983, pp 161-177
http://acl.ldc.upenn.edu/J/J83/J83-3003.pdf - Last Visit: 6.2004
In this paper meta-rules and a control structure under which they are invoked as a framework for processing ill-formed input are proposed. The LHS refers to an error encountered during normal parsing and the RHS relaxes the violation. Examples discussed in the paper include violated grammatical tests, omitted articles, homonyms, spelling/typographical errors, unknown words, violated selection restrictions, personification, and metonymy. Also the important distinction between "absolutely illformed" and "relatively illformed" is made. Note that the parser is based on an ATN which seems not so common nowadays.

Wolff, Dieter: *Neue Technologien und fremdsprachliches Lernen - Versuch einer*

Bestandsaufnahme. In: *DaF*. 35, 1998, pp 136-140 + 205-211
In diesem zweiteiligen Artikel stellt Wolff die Ideen vor, die später in Rüschoff und Wolff 1999 als Buch ausführlicher dargestellt wurden. Grundlage ist hier der Versuch einer Strukturierung des Bereichs CALL mit einem Fokus auf konstruktivistische Ansätze. Hervorzuheben ist hier der Versuch der Strukturierung von Computerprogrammen, die einerseits damit einer Bewertung unterzogen werden und andererseits dahingehend beurteilt werden, wie geeignet sie für konstruktivistische Ansichten sind.

AV

Handke, Jürgen ; Intemann, Frauke: *The Interactive Introduction to Linguistics*. Ismaning : Hueber, 1999 Version 1.10
CD-ROM for Win9x/NT. This CD-Rom which came out as Version 2.00 in 2000 presents a multi media based introduction to linguistics. Most of the core areas of linguistics from phonetics to semantics are covered on an introductory level. Therefore it is well suited for beginning students of linguistics. It uses the possibilities of multi media very well especially (as one would expect) in the area of phonetics and phonology. Also the chapter on the language of the world profits from the use of a CD-Rom in order to protray the different sounds of the languages. Note that this CD-Rom is recommended as a supplement to a introductory course.

Heringer, Hans Jürgen: *Aus Fehlern lernen*. Augsburg : Universität Augsburg, 1995
http://www.philhist.uni-augsburg.de/faecher/germanis/daf/forschung/fehler/index.html
CD-ROM for Win9x/NT. Diese CD-Rom enthält ein einfaches Konzept zur Präsentation von Fehlern aus einem annotierten Fehlerkorpus für das Deutsche. Die Aufgabe besteht in der Bestimmung des Fehlers durch den Lerner. Anschließend kann dann nachgeschlagen werden, ob der Fehler richtig erkannt wurde. Interessant an dieser CD-Rom ist die Tatsache, dass die Sätze alle authentisch sind, allerdings so korrigiert worden sind, dass jeder Satz nur einen Fehler enthält. Diese Datenbank lässt sich auch nach XML konvertieren und so für eine automatische Analyse nutzen.

Books

Atwell, Eric: *The Language Machine*. London : British Council, 1999
http://www.comp.leeds.ac.uk/eric/atwell99.pdf - Last Visit: 5.2004
In this booklet Atwell tries to foresee what the future will bring with regard to language technology and language learning. The leading question is whether language learning is still relevant if comprehensive translation systems have been developed in the future. As this book was written for the British Council it is written in a rather popular style. On the one hand the current state-of-the-art in speech and language technology is explained and on the other hand

the possibilities for future developments are discussed. The booklet also addresses the problems of CL in popular terms. Finally it also highlights the industry and research scenarios with a specific focus on the UK.

Bloom, Benjamin (Ed.): *Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook I, Cognitive Domain*. New York : Longman, 1956
This is THE classical work with respect to educational goals. This is still the major reference and to my knowledge no really better solution has been found. Even though it almost half a century old the classification can still be used. Of course most other aspects seem outdated nowadays.

Chomsky, Noam: . Dordrecht : Foris, 1981

Fechner, Jürgen (Ed.): *Neue Wege im computergestützten Frendsprachunterricht*. Berlin : Langenscheidt, 1994

Goerz, Günther ; Rollinger, Claus-Rainer ; Schneeberger, Josef (Ed.): *Handbuch der künstlichen Intelligenz*. 3. München : Oldenbourg, 2000
Diese Buch bietet die einzige mir bekannte deutschsprachige Übersicht über Künstliche Intelligenz. Es teilt sich in die folgenden Kapitel auf: Kognition, Neuronale Netze, Suche und Constraints, Wissensrepräsentation, Logik und automatisches Beweisen, Unsicheres und vages Wissen, Wissen über Raum und Zeit, Fallbasiertes Schließen und modellbasierte Systeme, Planen, Maschinelles Lernen und Data Mining, Sprachverarbeitung, Bildverstehen, Robotik und Software-Agenten. Die Themen werden aber sehr unterschiedlich angegangen. Manchmal sind Vorkenntnisse unabdingbar und manchmal werden auch grundlegende Aspekte ausführlich erläutert.

Greer, Jim (Ed.): *Student modelling: the key to individualized knowledge based instruction*. Berlin : Springer, 1994

Grüner, Margit ; Hassert, Timm: *Computer im Deutschunterricht*. München : Langenscheidt, 2000
Dieses Buch beschreibt in sehr einfacher Form die Möglichkeiten der

Higgins, John: *Computers and English language learning*. Oxford : Intellect, 1995

Holland, V. Melissa ; Kaplan, Jonathan D. ; Sams, Michelle R. (Ed.): *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995
As opposed to Yazdani et al. (1992) this book concentrates on contributions to the development of ICALL systems from North America. The book is divided into 4 sections: Text-Based Language Tutors and Learning Environments, Dialogue-Based Language Games,

Graphics-Based Language Tutors and Learning Environments, and Theoretical Issues in Language Tutor Design and Assesment. Notably the book contains descriptions of the ALLP project, the ALICE-chan system and the program 'Herr Kommissar'. There also a few articles from scientists who are not normally connected with NLP, such as one by Nina Garrett and one by Brian McWhinney.

Jager, Sake ; Nerbonne, John ; Van Essen, Arthur (Ed.): *Language teaching and language technology*. Lisse : Swets and Zeitlinger, 1998
This book grew out of a workshop held at the University of Groningen in April 1997 a contains selected papers from this workshop. The articles are divided into seven sections: Speaking (3 papers), vocabulary (3 papers), grammar (3 papers), reading, writing and testing (4 papers), distance learning (3 papers), users: models and studies (5 papers), and reflections and visions (2 papers). The length of the papers ranges from 3 pages to 15 pages. Some projects are presented in more than one paper, notably the Glosser project and the RECALL project with 3 papers each. What the book clearly demonstrates as opposed to e.g. Holland et al. (1995) is that here a some user studies are already included. I think this demonstrates the change from well designed systems but too slow to be used (ALLP) to systems like Glosser, which may contain less NLP technology but was actually used by language learners.

Jung, Udo O.H.: *An international bibliography of computer-assisted language learning with annotations in German*. 1. Edition. Frankfurt am Main : Lang, 1988

Kearsley, G.P. (Ed.): *Artificial Intelligence and Instruction*. Reading, MA : Addison-Wesley, 1987

Krüger-Thielmann, Karin: *Wissensbasierte Sprachlernsysteme*. Tübingen : Narr, 1992
Dieses ist die Dissertation von Krüger Thielmann, in der Ansätze zur Fehlererkennung mit Hilfe unterschiedlicher Parsingstrategien auf der Basis von PROLOG beschrieben werden. Gegenüber sehr umfassenden Parsern bechränkt sich dieser Ansatz auf die Beschreibung des Zahlensystems im Französischen.

Levy, Michael: *Computer-Assisted Language Learning - Context and Conceptualization*. Oxford : Clarendon Press, 1997
Michael Levy's book tries to interpret the development of CALL and also sheds some light on the history of ICALL. The historical development is then used to argue for new strategies in the development and use of CALL-software. Finally the results of a survey are presented in which language teachers presented their views on the 'context and conceptualization' of CALL. From this Levy draws the conclusion that some sort of CALL-theory is still missing and desirable.

Lieber, Gothild: *An international bibliography of computer-assisted language learning with annotations in German*. 2. Edition (Jung, 1988). Frankfurt am Main : Lang, 1993

Ludewig, Petra: *Korpusbasiertes Kollokationslernen - Computer-Assisted Language Learning als prototypisches Anwendungsszenario der Computerlinguistik* : Habilitationsschrift an der Universität Osnabrück, Fachbereich Sprach- und Literaturwissenschaften, 2003
In diesem Buch wird das System LogoTax beschrieben, das es einem Fremdsprachenlerner ermöglicht, sein persönliches elektronisches Wörterbuch anzulegen. Dabei wird er insbesondere bei der Eintragung von Kollokationen unterstützt, da hierzu Beispielsätze für Kollokationen aus Korpora extrahiert werden. Das Buch bietet einen zunächst einen Überblich über CALL bzw. ICALL. Anschließend werden die Grundlagen für die Behandlung von Kollokationen sowohl aus theoretisch-linguistischer als auch aus lexikographischer Perspektive erläutert. Schließlich erfolgen didaktische Überlegungen und eine Beschreibung des Systems.

Maas, Utz: *Grundzüge der deutschen Orthographie*. Tübingen : Niemeyer, 1992

Menzel, Wolfgang: *Modellbasierte Fehlerdiagnose in Sprachlehrsystemen*. Tübingen : Niemeyer, 1992
In diesem Buch beschreibt Wolfang Menzel sein Konzept einer modellbasierten Fehlerdiagnose. Dabei wird die Grammatik als ein Modell implementiert, in dem zum Parsing jede morphosyntaktische Eigenschaft eines Wortes ausspezifiziert wird und zur Kontrolle der Kongruenz und der Rektion mit jedem anderen relevanten Wort verknüpft wird. Daraus ergibt sich die Möglichkeit, umfassende Fehleranalysen und Diagnosen zu erreichen.

Mitton, Roger: *English Spelling and the Computer*. London : Longman, 1996

Ritter, Markus: *Computer und handlungsorientierter Unterricht*. Donauwoerth : Ludwig Auer, 1995

Rüschoff, Bernd ; Wolff, Dieter: *Fremdsprachenlernen in der Wissensgesellschaft: zum Einsatz der Neuen Technologien in Schule und Unterricht*. Ismaning : Hueber, 1999

Sama Nwana, Hyacinth (Ed.): *Mathematical intelligent learning environments*. Oxford : Intellect, 1993

Schulmeister, Rolf: *Grundlagen hypermedialer Lernsysteme: Theorie - Didaktik - Design*. München : Oldenbourg, 1997
http://www.izhd.uni-hamburg.de/paginae/Book/Frames/Start_FRAME.html - Last Visit: 5.2004
The URL is a link to the English Version of this book. Relevant is especially chapter 6 about

'Intelligent Tutoring: What is ITS?'. Schulmeister looks at these systems not from a technical position but from a didactic position. Therefore he focuses on the relation between theories of learning and the possibilities of ITSs. The first part gives an introduction into the field describing the main components of an ITS. Then a comparison with expert systems follows. After that he presents a few cases of systems where looks at how good the systems actually model psychological theories of learning and tries to demonstrate that almost all ITS-research so far has failed at creating a new learning theory-oriented approach in ITS.

Sleeman, D. ; Brown, J. S. (Ed.): *Intelligent tutoring systems*. London : Academic Press, 1982

Storch, Günther: *Deutsch als Fremdsprache. Eine Didaktik*. Stuttgart : UTB, 1999
Dieses Buch stellt eine allgemeine Einführung in die Didaktik von Deutsch als Fremdsprache dar. Von besonderem Interesse ist hier, dass gerade nicht ein ausgeprägter Konstruktivismus vertreten wird. Besondere Elemente sind eher Aspekte wie Interkulturelles Lernen etc.

Swartz, M. L. ; Yazdani, Masoud (Ed.): *Intelligent Tutoring Systems for Foreign Language Learning*. Berlin : Springer, 1992
This is the first major publication on ICALL-systems especially from Europe. Almost all somewhat larger projects are contained in this book.

Thome, Günther ; Thome, Dorothea (Ed.): *Computer im Deutschunterricht der Sekundarstufe*. Braunschweig : Westermann, 2000

Tomita, M.: *Efficient Parsing of NL: A Fast Algorithm for Practical Systems*. Boston, Ma : Kluwer, 1986

Yazdani, Masoud (Ed.): *Multilingual multimedia: bridging the language barrier with intelligent systems*. Oxford : Intellect, 1993

Incollection

Chanier, Thierry ; Pengelly, Michael ; Twidale, Michael ; Self, John: *Conceptual Modelling in Error Analysis in Computer-Assisted Language Learning Systems*. In: Swartz, M. L. ; Yazdani, M. (Ed.) *Intelligent Tutoring Systems for Foreign Language Learning*. Berlin : Springer, 1992, pp 125-150

deSmedt, William H.: *Herr Kommissar: An ICALL Conversation Simulator for Intermediate German*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995
In this text the system 'Herr Kommissar' is presented. Through dialogs with the system the user is in charge of solving crimes. The main task consists of interrogating the suspects in

order to identify the guilty person. The system has a knowledge management module and can therefore interpret the answers with respect to the knowlegde contained in the system. For example it can remind the user that a certain question has already been asked before. As the complete system is described the single modules are not described in much detail which makes it difficult to evaluate the concept of the program. This program was sold once for Apple-Computers but is apparently not available anymore.

Dokter, Duco ; Nerbonne, John ; Schurcks-Grozeva, Lily ; Smit, Petra: *Glosser-RuG: a User Study*. In: Jager, S. ; Nerbonne, J. ; Van Essen, A. (Ed.) *Language Teaching and Language Technology*. Lisse : Swets and Zeitlinger, 1998, pp 167-176
This article presents an overview of the Glosser system. The system is presented especially from a user's perspective. In the second part a small user study is presented. The outcome is not surprising as most users preferred the electronic version and of these most used the dictionary entries for reference. However it is one of the few user studies which are prepared using an ICALL-system which makes it more valuable.

Felshin, Sue: *The Athena Language Learning Project NLP System: A Multilingual System for Conversation-Based Language Learning*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995
This is one of the major references for the ALLP-project. It introduces all the relevant features of the system and notes especially the architecture with surprisingly complex modules.

Garrett, Nina: *ICALL and Second Language Acquisition*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995, pp 345-358

Hamburger, Henry: *Tutorials tools for language learning by two-medium dialogue*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995, pp 183-199
In this text the system FLUENT is presented. The system presents a microworld (a kitchen scene) in which the learner can either give orders to make things happen in the kitchen (e.g. put water on the stove) or the learner can describe the things she sees happening in the scene. The system is partly based on the modules developed in the ALLP project.

Handke, Jürgen: *Wizdom: A multi-purpose language tutoring system based on ai techniques*. In: Swartz, M. L. ; Yazdani, M. (Ed.) *Intelligent Tutoring Systems for Foreign Language Learning*. Berlin : Springer, 1992

Hu, Quian ; Hopkins, Jeff ; Phinney, Marianne: *Native English Writing Assistant - A CALL Product for English Reading and Writing*. In: Jager, S. ; Nerbonne, J. ; Van Essen, A. (Ed.) *Language teaching and language technology*. Lisse : Swets and Zeitlinger, 1998, pp 95-100

This short article introduces a commercial grammar- and spell-checker and the possible uses in a language learning scenario. The grammar checker uses a collection of rules retrieved from a learner corpus in order to check incorrect sentences in a text processor. Strangely though, the suggestions from the system for corrections are presented as an innovative application in CALL.

Levin, Lori S. ; Evans, David A.: *ALICE-chan: A Case Study in ICALL Theory and Practice*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995, pp 77-98
This article presents the system ALICE-chan, which also uses a parser to analyse learner input. It uses an LFG-like format. One important aspect the use of LFG's lexical mapping theory in order to achieve sufficient parsing results.

Murphy, Maureen ; Krüger, Anja ; Grieszl, Andrea: *RECALL - Providing an Individualized CALL Environment*. In: Jager, S. ; Nerbonne, J. ; Van Essen, A. (Ed.) *Language Teaching and Language Technology*. Lisse : Swets and Zeitlinger, 1998, pp 62-73
This paper (again) presents an overview of the RECALL-project and the resulting demonstrator. A number of components such as a parser including the diagnosis process and a tutoring module including a fairly detailed learner model are sketched. Due to the complexity of the system only little technical details are given.

Murray, Janet H.: *Lessons Learned from the Athena Language Learning Project*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995

Nerbonne, John ; Jager, Sake ; Van Essen, Arthur: *Language Teaching and Language Technology: An Introduction*. In: Jager, S. ; Nerbonne, J. ; Van Essen, A. (Ed.) *Language teaching and language technology*. Lisse : Swets and Zeitlinger, 1998, pp 1-10

Nerbonne, John: *Computer-Assisted Language Learning and Natural Language Processing*. In: Mitkov, Ruslan (Ed.) *Handbook of Computational Linguistics*. Oxford : Oxford University Press, 2002, pp 670-698
This article presents an overview including various subtopics in the field of ICALL. Interestingly the text also includes some notes about the question why NLP in CALL did not have the expected success so far as a "stable" technology.

Rypa, Marikka ; Feuerman, Ken: *CALLE: An Exploratory Environment for Foreign Language Learning*. In: Holland, V.M. ; Kaplan, J.D. ; Sams, M.R. (Ed.) *Intelligent Language Tutors*. Mahwah, NJ : Erlbaum, 1995, pp 55-76
The system described here is one which uses a parser but tries to teach the learner by

displaying the linguistics structures of authentic sentences to the learner. The idea is that the grammar theory LFG is quite close to traditional grammatical concepts and therefore can be used to make understanding of foreign texts easier and to teach grammatical structures.

Thome, Günther: *Experimente mit dem Computer im Grammatikunterricht der Sekundarstufe*. In: Thome, Günther ; Thome, Dorothea (Ed.) *Computer im Deutschunterricht der Sekundarstufe*. Braunschweig : Westermann, 2000, pp 78-87

Vogel, Carl ; Cooper, Robin: *Robust Chart Parsing with Mildly Inconsistent Feature Structures*. In: Schöter, Andreas ; Vogel, Carl (Ed.) *Nonclassical Feature Systems* : Edinburgh University, 1995, pp 197-216
This paper describes how to achieve robust parsing in a HPSG style grammar system. A certain type of unification is developed which leads to a 'bottom' as the value of a attribute in case of clashing values. This has the advantage that no error needs to be anticipated in any way. However no further trace of the error is kept in the feature structure and therefore no feedback can be given based on this. Additionally it is not clear how the system works with structure sharing which is a important part of HPSG.

Warschauer, Mark: *Computer-assisted language learning: An introduction*. In: Fotos, S. (Ed.) *Multimedia language teaching*. Tokyo : Logos International, 1996, pp 3-20
http://www.gse.uci.edu/markw/call.html - Last Visit: 5.2004
This article provides brief overview of how computers have been used and are being used for language teaching. It focuses not on a technical description of hardware and software, but rather on the pedagogical questions that teachers have considered in using computers in the classroom. The introduction follows a historical path from 'behaviouristic CALL' to the Internet. The appendix provides a rather comprehensive typology of CALL programs and a list of further CALL resources. Note that this text is from 1996 and therefore some aspect such as constructionist approaches to language learning are missing. Nevertheless the paper gives a good introduction and overview from a certain perspective.

Woolf, B. P.: *Theoretical Frontiers in Building a Machine Tutor*. In: Kearsley, G.P. (Ed.) *Artificial Intelligence and Instruction*. Reading, MA : Addison-Wesley, 1987, pp 229-267

Zock, Michael: *SWIM or SINK: the Problem of Communicating Thought*. In: Swartz, M. L. ; Yazdani, M. (Ed.) *Intelligent Tutoring Systems for Foreign Language Learning*. Berlin : Springer, 1992, pp 235-247
This article presents a new type of ICALL system as it supports a learner in creating meaningful utterances in a graphical user-interface. The learner is asked to put the elements of an utterance together by building a "graph-like" structure. The system than checks whether the nodes make sense according to the selection restrictions.

Inproceedings

Abel, Andrea ; Gamper, Johann ; Knapp, Judith ; Weber, Vanessa: *Evaluation of the Web-based Learners Dictionary ELDIT*. In: *Proceedings of the ED-MEDIA Conference*, 2003
In this text the authors present some data of an user evaluation of the ELDIT system. However the evaluation does not test whether the language learning with the system or the access of the system is better than with other electronic dictionaries. The evaluation merely finds out that users like the style of the presentation, that the users seldom use the help pages etc. Therefore from a didactic perspective this evaluation has little to offer.

Beck, Kathrin: *Ein Vokabeltrainer auf der Grundlage von GermaNet und MAPA (Mapping Architecture for People's Associations)*. In: *Proceedings des GermaNet-Workshops 2003*. Tübingen, 2003, pp 46-55
http://www.sfs.uni-tuebingen.de/lsd/GermaNet-Workshop/Mapa-Paper.pdf - Last Visit: 5.2004
Dieser Beitrag stellt ein in einem studentischen, interuniversitären Projekt entwickeltes Programm zur Konstruktion von 'Wissensnetzen' vor. Dabei soll der Benutzer Wörter und andere elektronische Daten aller Art verlinken können, um sein Wissen auf eine kognitiv adäquate Art und Weise zu repräsentieren. Als Beispielapplikation in diesem Projekt wurde ein Vokabeltrainer konzipiert, der die Daten von GermaNet als Netzstruktur darstellt, mit dessen Hilfe Lerner deutsche Wörter verstehen und lernen können. Vorgestellt werden unterschiedlich Möglichkeiten der Exploration des Wortschatzes und der Aufgabengenerierung. Allerdings handelt es hier tatsächlich nur um eine prototypische Implementierung, die nicht mit Lernern getestet wurde.

Bernedo, Gordon ; Elbers, Michael: *MAPA - A platform for collaborative, cognitively adequate knowledge mapping*. In: *Proceedings of the EuroCogSci'03*. Osnabrück, 2003, pp 372
In this poster the outline of the MAPA-project is sketched. Note that it is only a one-page abstract of the not yet finished project and contains mostly background ideas and a few hints of the implementation.

Borin, Lars: *Where will the Standards for Intelligent Computer-Assisted Language Learning Come from?*. In: *Proceedings of the LREC-Workshop on International Standards of Terminology and Language Resources Management*. Las Palmas, 2002, pp 61-68
An important trend in ICT-based language learning is that of standardization and resuability. Standards formats for all aspects of so-called 'instructional management systems' are rapidly gaining acceptance in the e-learning industry. In the paper it is discussed how ICALL applications can be related to the various standards proposals from 'traditional CALL' and

computational linguistics, basing the discussion on concrete experiences of the author from a number of ICALL-projects, where these standards are used or where their use has been contemplated. Note that only projects are presented a little more detailed, in which the author took part, which is therefore quite selective.

Foster, Jennifer: *A Unification Strategy for Parsing Agreement Errors*. In: Piliere, Catherine (Ed.) *Proceedings of the 5th ESSLLI 2000 Student Session*. Birmingham, 2000
In the paper a new kind of 'robust unification' is presented which takes two incosistent values at the root of an agreement violation and relates them using a relation which is called 'inconsistent identity'. The approach has the advantage that it acknowledges errors yet still produces consistent results. However it is argued that this robust unification is only appropriate for agreement errors and not for constraint violation errors which do not involve an agreement violation, which limits the approach to a certain extend.

Fouvry, Frederik: *Robust Unification for Linguistics*. In: *ROMAND 2000 1st workshop on RObust Methods in Analysis of Natural language Data*. Lausanne, 2000
In this paper, an approach that is based on a more abstract and general level of robust parsing, viz. the formalism is presented. More concretely, a typed attribute logic (i.e. HPSG) is used based on Carpenter, (1992). In a first step it is ensured that the type hierarchy has certain properties to make a meaningful robustness possible, i.e. containing information about what went wrong. Then, a distance measure based on the logic is introduced to distinguish between different analyses. This paper describes a similar approach as Fouvry (2003) as it is also part of his dissertation.

Fouvry, Frederik: *Constraint relaxation with weighted feature structures*. In: *Papers from IWPT2003, 8th International Workshop of Parsing Technologies*. Nancy, 2003
http://www.coli.uni-sb.de/~fouvry/publications/fouvry:03c.pdf - Last Visit: 5.2004
In this paper, a definition of unification of weighted feature structures designed to deal with constraint relaxation is presented. The application of phrase structure rules in a unification-based Natural Language Processing system is adapted such that inconsistent values do not lead to failure, but are penalised. These penalties are based on the signature and the shape of the feature structures. Note that this approach crucially depends on typed feature structures.

Gamper, Johann ; Knapp, Judith: *Adaptation in a Language Learning System*. In: *Online-Proceedings des 9. GI-Workshops: ABIS-Adaptivitaet und Benutzermodellierung in interaktiven Softwaresystemen*, 2001
http://www.kbs.uni-hannover.de/~henze/ABIS_Workshop2001/ABIS_2001.html
This paper describes the beginnings of the ELDIT-system, an electronic vocabulary acquisition system. The focus is on the adaptive features and outlines the authors' first ideas

about the use of adaptation technologies for vocabulary acquisition. As only first ideas are presented no 'substantial results' are presented yet. Therefore some options are mentioned but their effectiveness and usability remains unclear.

Hamburger, Henry: *Viewpoint Abstraction: a Key to Conversational Learning*. In: Appelo, L. ; de Jong, F. (Ed.) *7th Twente Workshop on Language Technology - Computer-Assisted Language Learning (TWLT7)*. Enschede : Universiteit Twente, 1994, pp 23-32

John, Roul Sebastian: *PROMISE: Steps towards Communicative English Language Teaching in an Interactive CALL System*. In: Appelo, L. ; de Jong, F. (Ed.) *7th Twente Workshop on Language Technology - Computer-Assisted Language Learning (TWLT7)*. Enschede : Universiteit Twente, 1994, pp 117-118

Kato, Tsuneaki: *Yet another Chart-Based Technique for Parsing Ill-Formed Input*. In: *Proc. 4th Conference on Applied Natural Language Processing (ANLP)*. Stuttgart, 1994, pp 107-112
http://acl.ldc.upenn.edu/A/A94/A94-1018.pdf
This paper takes up the approach from Mellish (1989) and tries to generalize from it to cover more errors and to make the recognition phase a little bit more straight forward. It also uses the same basic mechanism of two seperate parsing phases. However it does not rely on the complicated set of heuristics suggested by Mellish. Additionally Kato's approach uses a bi-directional parser in the first phase and a normal parser instead of a chart parser in the second phase. Most likely one can assume an admissable A* search for a plausible candidate. Note that only phrase structure rules are used and therefore some morphosyntactic information is missing from the results.

Kronenberg, Friedrich ; Krüger, Anja ; Ludewig, Petra: *Contextual Vocabulary Learning with CAVOL*. In: Appelo, L. ; de Jong, F. (Ed.) *7th Twente Workshop on Language Technology - Computer-Assisted Language Learning (TWLT7)*. Enschede : Universiteit Twente, 1994, pp 47-56
In this paper a program for vocabulary learning is presented which uses rather classical multiple choice exercises in general. However the selection of the items (distractors) is based on a dictionary structured according to lexical sematic relations 'a la' WordNet. Therefore the items in an exercise can be selected based on the semantic relation between them which should give a better learning curve. This same concept can be used to present exercises where a word with a certain sematic relation is asked for. Note however that the system was never evaluted with real learners.

Lee, Kong Joo ; Kweon, Cheol Jung ; Seo, Jungyun ; Kim, Gil Chang: *A Robust Parser Based on Syntactic Information*. In: *Proceedings of the 7th Conference of the European Chapter of*

the Association for Computational Linguistics (EACL). Dublin, 1995, pp 223-228
http://acl.ldc.upenn.edu/E/E95/E95-1031.pdf - Last Visit: 5.2004
A general algorithm for least-errors recognition, which is based only on syntactic information, was proposed by G. Lyon (1974) to deal with the extragrammaticality. This algorithm is extended in order to recover an extragrammatical sentence into a grammatical one in running text. The robust parser with recovery mechanism - an extended general algorithm for least-errors recognition - can be easily scaled up and modified because it utilizes only syntactic information. The style of parsing is based on Earley's algorithm. For each position in a sentence an error is hypothezised, which makes the approach difficult with respect to the number of chart items.

Mellish, Chris S.: *Some Chart-based Techniques for Parsing Ill-Formed Input*. In: *Proc. 27th Conference of the Association for Computational Linguistics (ACL)*, 1989, pp 102-109
http://acl.ldc.upenn.edu/P/P89/P89-1013.pdf - Last Visit: 6.2004
This paper presents a two phase, active chart parsing approach for parsing ill-formed sentences. In the first phase a bottom-up parser tries to fill the chart with as many items as possible. In the second phase a top-down approach tries to combine the items with the help of some abstract correction rules in order to build a complete phrase structure tree. This approach has been one of the most influential non-statistical concepts for robust parsing.

Menzel, Wolfgang: *Fehlerdiagnose und Feedback in einem Aussprachetrainer für den Fremdsprachenerwerb*. In: *Proc. 10. Arbeitstreffen der GI-Fachgruppe 'Intelligente Lehr- und Lernsysteme'*. Hamburg, 2000, pp 69-75
http://nats-www.informatik.uni-hamburg.de/%7Ewolfgang/papers/ills2000.ps.gz - Last Visit: 6.2004
This paper (in German!) reports about the modules in the ISLE-system for the analysis of spoken learner language. Several components do a phonetic diagnosis of learner language and produce error hypothesis for the learner. Some studies are presented which compare the analysis results with that of human raters. This is a rather short text as most of the components are described in various other publications.

Menzel, Wolfgang ; Schröder, Ingo: *Constraint-based Diagnosis for Intelligent Language Tutoring Systems*. In: *Proceedings of IT and KNOWS, XV. IFIP World Computer Congress*. Vienna/Budapest, 1998, pp 484-497
http://nats-www.informatik.uni-hamburg.de/%7Ewolfgang/papers/ifip98.ps.gz - Last Visit: 6.2004
This paper presents some results of applying a weighted constraint-based parsing approach in a ICALL scenario. As in principle every contraint can fail, any type of error can be found. However to make the system faster, constraints are weighted so that some constraint are hard,

i.e. they cannot fail. The basic concept of the parsing approach has been presented in a number of other publications. Note that this paper describes only a parser - a framework for actual language learning has not been developed.

Michaud, Lisa N. ; McCoy, Kathleen F.: *Evaluating a Model to DisambiguateNatural Language Parses on the Basis of User Language Proficiency*. In: *Proceedings of the 9th International Conference on User Modeling*. Heidelberg : Springer, 2003, pp 96-105
http://www.eecis.udel.edu/research/icicle/pubs/MichMcCo03.ps - Last Visit: 6.2004
This paper discusses the evaluation of an implemented user model in ICICLE, the instruction system for users writing in a second language. It is shown that in the task of disambiguating natural language parses, a blended model combining overlay techniques with user stereotyping representing typical linguistic acquisition sequences captures user individuality while supplementing incomplete information with stereotypic reasoning.

Murphy, Maureen ; McTear, Michael: *Learner Modelling for Intelligent CALL*. In: Jameson, Anthony ; Paris, Cecile ; Tasso, Carlo (Ed.) *User Modelling: Proceedings of the Sixth International Conference, UM97*. Wien : Springer, 1997, pp 301-312
This paper presents some result from the ReCALL project. The aim of this project was to work towards providing a more adequate and user-oriented interface for CALL. The adaptation was accomplished by developing a module that provides a parser-based analysis of the learner's response to the exercises of the program, a module that creates a model of the learner, and a module that controls the system's reactions to the learner's input and the structure of the materials offered to the learner.

Reuer, Veit: *Dialogue Processing in a CALL-System*. In: *Proc. 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. Bergen, 1999, pp 253-256
http://acl.ldc.upenn.edu/E/E99/E99-1037.pdf - Last Visit: 6.2004
In this paper two major knowledge bases for usage in an ICALL dialogue system are presented. A discourse grammar structures the dialogue elements (or dialogue acts) as possible parts of a dialogue and the dialogue knowledge base provides the possible contents of dialogues. A FSA based on the discourse grammar determines the possible moves which the dialogue might take. It is argued that on the one hand this concept is restricted enough to allow for (relatively) easy maintenance as well as expansion and on the other hand it is advanced enough to allow for simulated complex dialogues. Note that these modules have been used in other scenarios as well and represent nothing new as such.

Reuer, Veit: *Error-Recognition and Parsing of Syntactically Mildly Ill-formed Natural Language*. In: *Proc. LFG00 Conference*. Berkeley : CSLI Publications, 2000, pp 215-225

Schneider, David ; McCoy, Kathleen F.: *Recognizing Syntactic Errors in the Writing of Second Language Learners*. In: *Proc. 17th Int. Conference on Computational Linguistics (COLING)*. Montreal, 1998
http://acl.ldc.upenn.edu/P/P98/P98-2196.pdf - Last Visit: 6.2004
This paper reports on the recognition component of an intelligent tutoring system that is designed to help foreign language speakers (signers of American Sign Language (ASL)) learn standard English. The method of capturing ungrammaticalities involves using mal-rules, even though the straightforward addition of some mal-rules causes significant performance problems with the parser. The grammar is evaluated on its ability to correctly diagnose agreement problems in 'only' 70 sentences produced by ASL native speakers. This is one of the few ICALL-papers ever having appeared in a publication of the ACL.

Schröder, Hartmut: *Evaluierungskriterien für multimediale Lernprogramme. Ein Raster für die Praxis*, 1996
http://www.sw2.euv-frankfurt-o.de/Publikationen/FsU/frame/ - Last Visit: 5.2004

Schwind, Camilla: *Sensitive Parsing: Error Analysis and Explanation in an Intelligent Language Tutoring System*. In: *Proc. 12th Int. Conference on Computational Linguistics (COLING)*. Budapest, 1988, pp 608-613
http://acl.ldc.upenn.edu/C/C88/C88-2127.pdf - Last Visit: 5.2004
The paper present a framework for dealing with errors in natural language sentences within a proposed ICALL system. The idea is to use a feature grammar and to analyse errors as being sentences where features have other values than those they should have. By using a feature grammar it is possible to describe various types of errors (agreement, syntactic and semantic errors) in a uniform framework. However several different types of feature need to be considered for the different types of errors. Agreement feature are not a problem, but 'syntactic errors' do need additional rules in the grammar to anticipate the errors where the error itself is identified via clashing features. This also goes for semantic errors.

Schwind, Camilla: *Error analysis and explanation in knowledge based language tutoring*. In: Appelo, L. ; de Jong, F. (Ed.) *7th Twente Workshop on Language Technology - Computer-Assisted Language Learning (TWLT7)*. Enschede : Universiteit Twente, 1994, pp 77-92

Self, John A.: *Knowledge, Belief and User Modelling*. In: O'Shea, Tim ; Sgurev, Vasil (Ed.) *Artificial Intelligence III: Methodology, Systems, Applications - Proceedings of the Third International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA '88)*. North-Holland, 1988, pp 3-9

Thomann, Johannes: *LFG as a Pedagogical Grammar*. In: *Proceedings of LFG02*. Athens,

2002, pp 366-372
http://cslipublications.stanford.edu/LFG/7/lfg02thomann.pdf - Last Visit: 5.2004
The paper describes a presentation format of grammatical information for language teaching. C-structures and f-structures are represented as graphical annotation to a text. It is advocated that LFG-like concepts are preferable to traditional grammar rules. Note that this paper appeared as a poster is therefore only 3 real pages short, i.e. no indepth discussion is included.

Virvou, Maria ; Tsiriga, Victoria: *Web Passive Voice Tutor: An Intelligent Computer Assisted Language Learning System over the WWW*. In: Okamoto, T. ; Hartley, R. ; Kinshuk ; Klus, J.P. (Ed.) *Proceedings IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges*, 2001, pp 131-134
http://thalis.cs.unipi.gr/~vtsir/Virvou_Tsiriga_ICALT2001.pdf - Last Visit: 5.2004
In this paper the so called Web Passive Voice Tutor (Web PVT), an adaptive web-based Intelligent Computer Assisted Language Learning (ICALL) program is described that is aimed at teaching non-native speakers the passive voice of the English language. Web PVT incorporates techniques from ITS and Adaptive Hypermedia technologies to provide students with individualized instruction and feedback. The system uses a combination of stereotypes and the overlay technique for the initialisation of the student model, which is then refined by observing the student while working with the system. The resulting student model is used for the annotation of the links to topics presented to the student. However the techniques used in the system are only mentioned in the text and not explained. Therefore only little information about the system is contained in the text as it is only described very briefly.

Zock, Michael: *Sorry, but what was your name again, or, how to overcome the tip-of-the tongue problem with the help of a computer?*. In: *Proceedings of the COLING-Workshop on Building and Using Semantic Networks*. Taipeh, 2002
The paper presents a concept for the structuring of an electronic dictionary following especially psycholinguistic facts about access to the mental lexicon. Work done by psychologists reveals that people being in this so called tip-of-the-tongue state (TOT) know a lot about the word : meaning, number of syllables, origine, etc. Speakers are generally able to recognize the word, and if they produce an erroneous word, that token shares many things with the target word (initial/final letter/phoneme, part of speech, semantic field, etc.). Three methods (access by form, access by meaning and a combination of both) that assist the speaker/writer by revealing the word that's on his/her mind (tongue/pen) are presented, the first one being implemented.

Internet

Wazel, Gerhard: *Evaluationskriterien für sprachliche Multimedia-Software*, 2000
http://www.iik.com/theorie/theoretisches/kriterien.html - Last Visit: 6.2004

Hierbei handelt es sich um einen AUszug aus einem umfangreicheren Papier. Der Kokus liegt einerseits auf allgemeinen Kriterien zur Bewertung von CALL-Software. Andererseits werden aber auch deutlich Empfehlungen gegeben, welche Fähigkeiten ein Programm besitzen sollte, ohne allerdings auf die technischen Leistungen einzugehen. Verschiedene Bereiche werden in geschachtelten Listen detailliert dargestellt.

Phdthesis

Heift, Trude: *Designed Intelligence: A Language Teacher Model*. Burnaby, B.C., 1998
http://www.sfu.ca/langlab/trude/pubs.html - Last Visit: 6.2004
This is the dissertation which presents the groundwork for the system 'German Tutor', later known as 'E-Tutor' which has been presented in numerous publications. It uses the grammar formalism HPSG in order to analyse erroneous sentences in a more or less non-anticipating way. The system is only able to detect very specific constituent errors, i.e. errors in linear precedence. Apart from recognizing errors it also has a module for disambiguating sentence readings and a student module for 'informed' feedback.

Reuer, Veit: *PromisD - Ein Analyseverfahren zur antizipationsfreien Erkennung und Erklärung von grammatischen Fehlern in Sprachlehrsystemen*, 2003
In dieser Dissertation wird das System PromisD beschrieben, das auf einigen Ideen des Studienprojekts Promise (Bauer 1994) aufbaut. Zunächst findet eine Einordnung des Systems innerhalb des Gebietes CALL und die Behandlung von fehlerhaften Äußerungen von Fremdsprachenlernern statt. Anschließend wird das von Reuer entwickelte System präsentiert, wobei der Schwerpunkt auf der Vorstellung des Fehler-sensitiven Parsing-Algorithmus liegt. Es wird versucht, jeweils ein Verfahren zum sensitiven Verfahren für die aus der LFG bekannten F- und C-Strukturen zu finden.

Proceedings

Appelo, Lisette ; de Jong, Franciska (Ed.): *7th Twente Workshop on Language Technology - Computer-Assisted Language Learning (TWLT7)*. Enschede : Universiteit Twente, 1994
These are the proceedings of the 7th Twente Workshop on Language Technology which follows a certain topic every year. In 1994 the theme was "Advances in Computer-Assisted Language Learning". Some major systems were presented in this volume such as Zock's "SWIM OR SINK", Hamburger's "FLUENT" and Schwind's algorithm for error recognition. On the other hand the articles contained are only extended versions of talk given at the conference and are therefore not very long.

Carlson, R. ; Dunger, C. ; Granstrom, B. ; Oster, A. (Ed.): *STiLL - Speech Technology in Language Learning: ESCA workshop*. Stockholm : ESCA, 1998

Techreport

Althaus, Nadja ; Beck, Kathrin ; Bennöhr, Jasmine ; Bernedo, Gordon ; Boeck, Manuel ; Elbers, Michael ; Kugel, Felix ; Scherbaum, Stefan ; Widdra, Tobias ; Wissmann, Jens: *Abschlussbericht des Studentenprojekts MAPA im Masterprogramm Cognitive Science*. Osnabrück, 2003
In the MAPA study project, students from the Universities of Osnabrück, Tübingen and Bochum participated in a joint effort to develop a framework that allows the mapping of knowledge in a cognitively adequate way. The approach connects to existing techniques of mind mapping, concept mapping and the like. A major assumption is that the internal mental representation of knowledge is networklike and that therefore its externalisation must be done accordingly. Due to the complexity of knowledge however, the external representations consist only of cues to knowledge rather than a representation of knowledge itself. The report reports on the foundations of mind mapping, presents the implemented system and has an extra chapter about an vocabulary trainer integrated into the framework (Beck 2003).

Bauer, Pascal ; John, Roul S. ; Kronenberg, Friedrich ; Krüger, Anja ; Menzel, Andre ; Reuer, Veit ; Unsöld, Robert: *Abschlussbericht PROMISE Studienprojekt*. Osnabrück, 1994
PROMISE is an acronym for "PROjekt Mediengestütztes Interaktives Sprachelnlernen Englisch" (i.e. media supported interactive language learning - English). In PROMISE, the communicative approach known from language pedagogy is realized by putting the learner in an adventure-game-like setting where authentic and purposeful language use is possible. Within a situational frame (e.g. a road accident) the student is guided through a series of dialogue exercises with "free" learner input in which the student "talks" with a simulated dialogue partner in the foreign language and gets meaningful responses from this partner. In order to realize this a number of modules need to interact generating the appropriate feedback especially with respect to language errors the learner makes. An LFG-parser is included as well as a module for handling the dialogue. Furthermore a world knowledge database allows among others the special treatment of spatial expressions. A language learner can therefore be supported when learning the usage of these patterns.

Bongardt, Gerd ; Dalinghaus, Klaus ; Dittman, Hendrik ; Huber, Michael ; Krumeich, Alexander ; Peylo, Christoph ; Rehm, Georg ; Rother, Andreas ; Teiken, Wilfried: *Semf-Abschlussbericht*. Osnabrück, 1996
http://www.cogsci.uni-osnabrueck.de/~semf/ - Last Visit: 5.2004
In diesem Bericht werden die Ergebnisse des SEMF-Projekts vorgestellt. In diesem Projekt wurde versucht, Fehlleistungen eines Sprachlerners im semantischen Bereich zu erkennen und dazu Feedback zu bieten. Grundlage dafür ist eine so genannte Blocksworld, in der der Lerner Zustände beschreiben kann. Auf der Grundlage des vom System berechneten Zustandes der

Welt kann eine entsprechende Rückmeldung gegeben werden.

Erpenbeck, Arno ; Koch, Britta ; Kummer, Norman ; Reuter, Philip ; Tschorn, Patrick ; Wagner, Joachim: *KOKS-Abschlussbericht*. Osnabrück, 2001
http://www.cogsci.uni-osnabrueck.de/~koks/
In der ersten Phase wurden zur Indentifizierung der Kollokationen mit Hilfe eines Taggers und eines bilingualen Lexikons zunächst die Sätze alignet. Basierend auf einer statistischen Analyse der hochfrequenten POS-Abfolgen in den Korpora wurden Chunking-Regeln entwickelt, die in einer zweiten Phase ein Phrasen-Alignment ermöglichen. Darauf aufbauend wurde in einer dritten Phase das bilinguale Lexikon herangezogen, um Phrasen zu identifizieren, deren alignetes Pendant nicht durch die Übersetzungen im Lexikon abgedeckt wird. Damit wurde ein Verfahren zur Identifikation von Kollokationen entwickelt, das sich besonders durch die Berücksichtigung linguistischer Kriterien in der Definition von Kollokationen auszeichnet.

Krüger, Anja ; Dittmann, Hendrik ; Murphy, Maureen: *Grammar Based Error Diagnosis in CALL*, 1997
Informatics Research Reports - University of Ulster. This report is probably only available by contacting one of the authors. This report presents the approach taken in the ReCALL project, in which a large coverage parser is extended to include rules for the semantic analysis of certain constructions. It by no means a general approach as every information has to be entered into the grammar the same way it has to be added to a knowledge base. However for certain situations this may be a feasible approach.

Sagvall Hein, Anna: *A Grammar Checking Module for Swedish (SCARRIE Deliverable 6.6.3)*. Uppsala, 1998
http://www.ling.uu.se/wp/wp12c.pdf - Last Visit: 6.2004
This report from the SCARRIE project presents the approach taken for grammar checking. It uses a bottom-up chart-parser (the Uppsala Chart Processor (UCP)) to generate as much chart items as the grammar allows. A second module then scans the chart with the help of error-rules in order to detect grammar errors. The system is not evaluated against a corpus but the report includes a detailed description of the errors covered.

Glossary

CALL	***C***omputer-***A***ssisted ***L***anguage ***L***earning: The field of applying information and communication technology in a language learning/teaching situation
CL	***C***omputational ***L***inguistics
ICALL	***I***ntelligent ***C***omputer-***A***ssisted ***L***anguage ***L***earning: The field of applying information and communication technology in a language learning/teaching situation and the application includes methods of computational linguistics.
ITS	***I***ntelligent ***T***utoring ***S***ystem: A learning technology system that dynamically adapts learning content to a learner's specific objectives, needs, and preferences by using its expertise in instructional methods and the subject to be taught. (A Definition from the "Learning Technology Standards Committee Glossary Working Group P1484.3" of the IEEE)
Regular Expression	A regular expression is a formal method for specifying search and replace operations on symbols.
WWW	The World Wide Web is a part of the Internet.