

## 5.10 Sprachlehr- und -lernsysteme

*Petra Ludewig*

Versuche, mit dem Computer (Fremd-)Sprachen zu lernen bzw. zu lehren, reichen bis weit in die 60er Jahre zurück und erfolgten zunächst auf Großrechnern. Aufgrund der zunehmenden Verfügbarkeit und Leistungsfähigkeit von PCs in Privathaushalten – insbesondere wegen ihrer wachsenden Multimediafähigkeit – und durch die steigende Relevanz von Fremdsprachenkenntnissen in Beruf und Freizeit hat die Verbreitung elektronischer Sprachlernhilfen in den 90er Jahren enorm zugenommen. Es werden nicht nur Vokabel-, Rechtschreib-, Aussprache- und Grammatiktrainer angeboten, sondern auch landeskundlich ausgerichtete Software, Adventurespiele sowie Schreibhilfen, Konkordanzprogramme für Korpusrecherchen und anderes mehr. Als Lernpartner werden dem Computer die Vorzüge eines a) zeitlich flexibel verfügbaren, b) sich dem individuellen Arbeitstempo und Kenntnisstand des Lerners anpassenden und c) objektiven Betreuers zugesprochen, der d) ein schnelles Feedback bietet, e) Lernprotokolle anfertigen kann und f) menschlichen Lernpartnern monotone Tätigkeiten abnimmt.

Aus der Sicht der Fremdsprachendidaktik geben *Rüschhoff und Wolff* 1999 einen ausführlichen Überblick über das breite Spektrum aktuell auf dem Markt erhältlicher Programme. Und eine ausgiebige Darstellung der historischen Entwicklung von Computer-Aided Language Learning (**CALL**) kann *Levy* 1997 entnommen werden. Ziel des vorliegenden Beitrags ist es, die Möglichkeiten und Grenzen von CALL-Systemen aus der Sicht von NLP-Technologie zu erläutern (s.a. *Nerbonne* 2003; *Jager et al.* 1998). Dabei bleiben solche Sprachlernprogramme unberücksichtigt, die sich auf digitale Videotechnik und internetbasierte Komponenten wie e-Mail, Chat, News-Groups und WWW-Browser beschränken, weil sie für computerlinguistische Betrachtungen uninteressant sind. Dennoch hat die oben angedeutete Vielseitigkeit von CALL als Anwendungsbereich zur Folge, dass sich im Grunde aus fast allen Bereichen der CL für das computergestützte Sprachenlernen Nutzen ziehen lässt. Für die weiteren Überlegungen sind insbesondere zwei Unterscheidungen von Bedeutung, die von CALL und ICALL und die von tutoriellen und toolartigen Systemen.

CALL stellt insofern keinen Teilbereich der Computerlinguistik dar, als ein Großteil, um nicht zu sagen der überwiegende Teil, der gemeinhin verfügbaren CALL-Software über keine NLP-Module verfügt. Um unter den CALL-Programmen diejenigen, die auf CL&KI-Technologien basieren, von den übrigen Sprachlernprogrammen abzugrenzen, werden sie gerne mit **ICALL** – Intelligent CALL – tituiert. Damit bezieht sich CALL im weiteren Sinne auf die Gesamtmenge aller Sprachlernprogramme und im engeren Sinne auf diejenigen Sprachlernprogramme, die auf CL&KI-Technologien verzichten.

**Tutorielle Systeme** gehen von einem fest umreißbaren Kanon zu lernender (sprachlicher) Inhalte aus. Diese Programme dienen vorrangig dazu, dem Lerner Übungsmöglichkeiten zu zuvor festgelegten Inhalten zu bieten. Neben derartigen tutoriellen Systemen gewinnen **toolartige Systeme** bzw. Lernwerk-

zeuge zunehmend an Bedeutung. Sie halten den ursprünglichen (sprachlichen) Lerngegenstand eher variabel und machen die Lerntechniken selbst zu einem zusätzlichen Lerngegenstand. Die Unterteilung in tutorielle und toolartige Systeme korrespondiert weitgehend mit der didaktischen Unterscheidung von Lernen durch Instruktion und Lernen durch Konstruktion (*Rüschhoff und Wolff 1999*).

### 5.10.1 Herkömmliche Sprachlehrprogramme

Kommerzielle CALL-Programme bieten in der Regel multiple-choice-Aufgaben, Lückentexte und solche Aufgaben an, bei denen der Lerner nur einzelne Wörter einzugeben hat, wie z.B. bei Vokabelgleichungen. Umfangreichere Lernereingaben können mit herkömmlichen Programmierverfahren nicht angemessen auf Korrektheit überprüft werden. Die Anzahl alternativer korrekter Eingaben und erst recht das Spektrum fehlerhafter Eingaben wachsen mit der Länge der zugelassenen Antworten, so dass ein explizites Auflisten aller potenziellen Lernereingaben inklusive der zugehörigen Rückmeldung an den Lerner unpraktikabel wird. Dies ist folgeschwer: Lerner reagieren extrem sensibel, wenn eine an sich korrekte Antwort vom System als falsch abgetan wird, und erwarten im Falle von Fehlern auch Hinweise auf die Art des Fehlers. Konkret bedeutet das: Als Übersetzung von *angry* darf nicht nur *wütend* akzeptiert werden, sondern auch *zornig* muss als richtig anerkannt werden. Auch die Beurteilungen *falsch* und *korrekt* reichen nicht aus. Eine falsch geschriebene Vokabel ist anders zu bewerten als eine "nicht gewusste" Vokabel. Darüber hinaus sind – selbst bei falsch geschriebenen Vokabeln – Performanz- von Kompetenzfehlern zu unterscheiden. Wenn z.B. in Antworten häufiger an Stelle von "n" ein "m" steht, weil die entsprechenden Tasten nebeneinander liegen, ist ein Tippfehler anzunehmen. Wenn beim Englischtraining jedoch statt eines "u" immer wieder ein "a" geschrieben wird, weil die Aussprache dieses deutschen Lernern in Wörtern wie *sunny* suggeriert, so ist dies als Rechtschreibfehler zu bewerten (vgl. Unterkapitel 5.1 über Korrekturprogramme), und ein Exkurs in eine entsprechende Rechtschreibübung anzubieten. Eine zu grobe Fehlerklassifikation führt zu einer oberflächlichen Fehlerstatistik, so dass der aktuelle Kenntnisstand des Lerners nur unzureichend modelliert und in die Auswahl weiterer Übungen einbezogen werden kann.

Zugegebenermaßen können beim Sprachenlernen immer wieder Themenbereiche abgesteckt werden, bei denen eine Verarbeitung komplexer Eingaben, etwa ganzer Sätze, verzichtbar ist. Darüber hinaus ist die Beherrschung einer fremden Sprache jedoch weitaus mehr als die Produktion einzelner Wörter. Der Lernende muss auch in die Lage versetzt werden, komplexe Äußerungen zu produzieren. Die Defizite der gegenwärtigen kommerziellen Sprachtutoren, bei denen NLP-Technologie ansetzen sollte, liegen also vornehmlich in den Bereichen:

- Fehlerbehandlung ,
- Lernermodellierung (vgl. Abschnitt 3.6.5 zur Benutzermodellierung) und
- Verarbeitung komplexer Eingaben.

### 5.10.2 ICALL-Systeme

Wie internationale Workshops und Tagungen – z.B. TWLT7 (*Appelo und de Jong* 1994), TALC (*Wichmann et al.* 1997) und EUROCALL (*ReCALL, Schulze et al.* 1999) – zeigen, wenden sich einige Computerlinguisten zwar in ihren thematischen Ausrichtungen der ICALL-Domäne zu, verzichten aber letztendlich auf die Realisierung eines elaborierten und didaktisch fundierten Sprachlehrszenarios, da die sprachdidaktische Einbettung erfahrungsgemäß sehr arbeitsaufwändig ist und den eigentlichen Kompetenzbereich des Computerlinguisten sprengt. Diese Vorgehensweise gilt auf Tool-Ebene z.B. für das DEFI-Projekt (*Michiels* 2001), das in den Bereich der Lesehilfen fällt und den Zugriff auf Übersetzungen von Wörtern und Mehrwortausdrücken unterstützt. Im tutoriellen Bereich widmen sich etwa *Menzel* 1992 und *Schwind* 1994 erweiterten Parsingtechniken zur syntaktischen **Fehlerdiagnose**. Ursprünglich dient das Parsing der Identifikation und Strukturbeschreibung grammatisch korrekter Sätze. In tutoriellen ICALL-Systemen geht es aber gerade um die Identifikation und Klassifikation von Fehlern und damit um die detaillierte Beschreibung grammatisch inkorrektur Sätze. Hier reicht noch nicht einmal wie bei anderen Anwendungskontexten (etwa natürlichsprachlichen Dialogsystemen, vgl. Unterkapitel 5.9) ein fehlerrobustes Parsing aus, bei dem Sätze fehlertolerant, d.h. die Fehler ausgleichend bzw. ignorierend, analysiert werden. Vielmehr ist ein fehlersensitives Parsing erforderlich, das Strukturbeschreibungen erzeugt, auf deren Grundlage Fehlermeldungen an den Benutzer zurückgegeben werden können. Die formalen Fehlerspezifikationen des fehlersensitiven Parsers sind dann in für den Sprachenlerner verständliche Fehlermeldungen umzusetzen.

#### Generische Beschreibung von ICALL-Systemen

Aus der Sicht sprachorientierter KI ist ein ICALL-System in der Regel ein Spezialfall eines Intelligenten Tutoriellen Systems (**ITS**)<sup>4</sup>, das über eine Wissenskomponente (Wissen über den zu vermittelnden Gegenstandsbereich, hier Sprachwissen), ein Didaktikmodell (Curriculum, Übungssammlung) und ein Lernermodell (Kenntnisstand und Lerngewohnheiten des Lerners) verfügt, um die Kompetenzen eines menschlichen Tutors simulieren zu können.

Gemäß der Architekturskizze in Abbildung 5.17 wird die Lernereingabe vom Tutormodul an die Sprachkomponente, das sogenannte Diagnosemodul, weitergereicht. Um dem Lerner angemessene Rückmeldungen zu geben und per Rückschluss auf dessen aktuellen Kenntnisstand weitere geeignete Übungen anzubieten, gilt es zu klären, um welchen Fehlertyp es sich handelt (ein orthographischer, morphologischer, syntaktischer oder lexikalischer Fehler), und gegen welche sprachliche Regel genau verstoßen wurde. Zwecks Fehlerdiagnose kann die Analysekomponente z.B. auf eine Fehlergrammatik zurückgreifen, welche Regeln beinhaltet, die fehlerhafte Äußerungen charakterisieren. Die Verwendung von Fehlerregeln stellt gewissermaßen ein Pendant zur erfahrungsbasierten Fehlerdiagnose menschlicher Tutoren dar.

<sup>4</sup>Für intelligente Sprachtools ist eine generische Beschreibung nicht möglich.

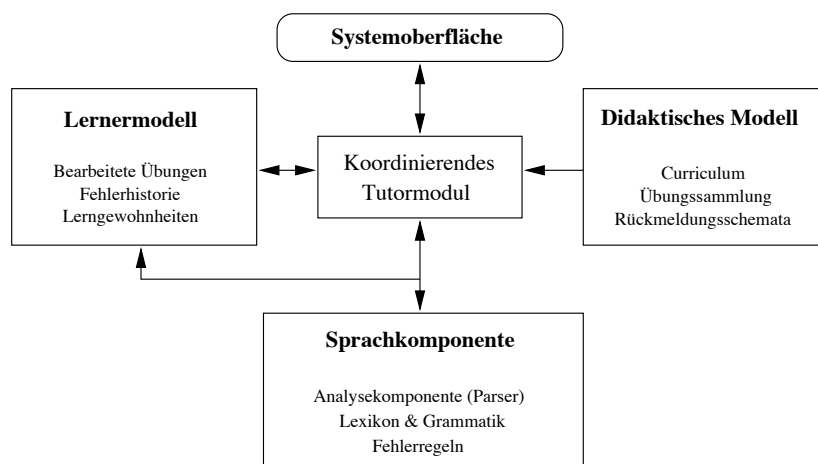


Abbildung 5.17: Architekturskizze eines ITS für ICALL

Ein Vorteil von Fehlerregeln ist, dass sie ein sehr genaues Fehlerfeedback erlauben, ein Nachteil, dass die verschiedenen Fehlertypen explizit vorweggenommen werden müssen. *Menzel 1992* und *Reuer 2000* bemühen sich aus diesem Grund um eine antizipationsfreie Fehlerdiagnose, bei der auf eine explizite Vorwegnahme von Fehlern verzichtet wird und stattdessen einzelne Bedingungen von Grammatikregeln zurückgenommen werden, um den Fehler zu lokalisieren. Dieses Verfahren erweist sich allerdings als sehr zeitaufwändig.

Unabhängig davon, für welche Variante der Fehlerdiagnose man sich entscheidet, tritt das Problem auf, dass zu einer fehlerhaften Lernereingabe oftmals konkurrierende Fehlerursachen angenommen werden können. Wo genau liegt z.B. der Fehler in dem Satz *Der Student sehen die Kommilitonen?* Liegt ein syntaktischer Fehler in Form eines Verstoßes gegen die Numeruskongruenz von Verb und Subjekt vor (wobei genau genommen unklar ist, ob das Verb oder das Subjekt zu modifizieren ist) oder ein morphologischer Fehler, weil der Lerner die Verbflexion nicht beherrscht? Wenn keine Kontextinformation nutzbar ist, macht es in solchen Fällen Sinn, eine Anfrage an die Fehlerhistorie im Lernermodell zu stellen, welcher der in Frage kommenden Regelverstöße für den betreffenden Lerner am wahrscheinlichsten ist. Gleichzeitig ist das Lernermodell natürlich in Hinblick auf den gerade identifizierten Fehler zu aktualisieren. Die Beziehung zwischen Diagnose- und Lernermodell sollte daher beidseitig sein. Für die Auswahl neuer Übungen muss das Tutormodul die Lernhistorie und das Curriculum mit den Lehrplänen berücksichtigen.

### Exemplarische ICALL-Prototypen

Auf dem Markt gibt es bis dato nur wenige ICALL-Systeme. Eine der wenigen Ausnahmen in Hinblick auf tutorielle Systeme ist das von Comet/Cornelsen vertriebene Programm WIZDOM (*Handke 1992*), das die Lernereingaben aber

ausschließlich auf der Grundlage von lexikalischem und morphologischem Wissen analysiert. Eine syntaktische Analyse ganzer Sätze findet nicht statt. Ein Beispiel für ein kommerzielles toolartiges System ist das von Oxford University Press vertriebene Konkordanzprogramm WordSmith, das unter anderem über eine Funktion verfügt, mit der für einen Text und dessen Übersetzung Korrespondenzen im Sinne eines Satzalignments bereitgestellt werden können (vgl. Abschnitt 4.1 zu Textkorpora sowie *Kettemann und Marko* 2002; *Wichmann et al.* 1997). Diese Funktion hat zur Folge, dass der Lerner sich zu einem Belegsatz auch dessen Übersetzung anzeigen lassen kann, was das Satzverständnis erleichtert, und dass sich außerdem fremdsprachliche Belege auch durch Vorgabe muttersprachlicher Suchraster abrufen lassen.

**Tutorielle Systeme:** Die Adventure-Systeme „Spion“ (*Sanders und Sanders* 1995) und „Herr Kommissar“ (*DeSmedt* 1995) sowie das Simulationssystem PROMISE (PROjekt Mediengestütztes Interaktives Sprachenlernen – Englisch) (*Krüger* 1996) sind dialogbasierte Lernprogramme und dem kommunikations- und handlungsorientierten Fremdsprachenlernen verpflichtet. Demzufolge wird die kommunikative Kompetenz des Lerners mit Hilfe situativ eingebetteter Rollenspiele gefördert, in denen der Lerner bestimmte sprachliche Mittel einsetzen muss, um bestimmte Handlungsabsichten zu verwirklichen.

Zu diesem Zweck versuchen diese Systeme – wenn auch nur in schriftlicher und nicht in akustischer Form – möglichst realistische Dialoge zwischen dem Lerner und simulierten Dialogpartnern zu gewährleisten. Dies setzt voraus, dass die Systeme die Eingaben des Lerners quasi verstehen. Dazu werden die Lernereingaben unter Einbeziehung von sprachlichem, konzeptuellem und situativem Wissen syntaktisch und semantisch analysiert (vgl. die Unterkapitel 3.4 und 3.5). In „Spion“ wird eine Kategorialgrammatik verwendet, in „Der Kommissar“ eine Kasusgrammatik und in PROMISE aufgrund ihrer Nähe zu Schülergrammatiken eine LFG-Grammatik. Um die von der NLP-Technologie zu bewältigende Aufgabe zu erleichtern, versuchen diese Systeme, die zu erwartenden Lernereingaben über das situative Szenario mit seinen begrenzten Objekten und Aktionen auf allen Ebenen (also z.B. lexikalisch, syntaktisch und inhaltlich) auf natürliche Weise einzuschränken<sup>5</sup>. So sind die Lernereingaben in „Spion“ aus technologischen Gründen auf Imperative begrenzt; dies erscheint jedoch für den Lerner natürlich und plausibel, da er den Spion in seinem Verhalten dirigieren muss.

Es handelt sich bei allen drei Systemen um hochgradig spezialisierte und strukturierte Dialogtypen. Sie zeichnen sich dadurch aus, dass sie einen hohen Anteil formelhafter Wendungen beinhalten und das Rollenverhalten der Gesprächspartner relativ stark festgelegt ist. Zum einen sind das rollenkonforme Verhalten und der Gebrauch einschlägiger Wendungen in derartigen Situationen Lerngegenstand. Zum anderen ist gerade das Schematische dieser Dialoge auch dafür verantwortlich, dass sich derartige Dialoge recht gut bewerkstelligen lassen. In PROMISE geschieht dies mit Hilfe von Gesprächsskripten, welche die Standardvarianten der Dialogverläufe in Form von Übergangsnetzwerken beschreiben.

<sup>5</sup>Eine alternative Strategie, die Anforderungen an die zugrundeliegende NLP-Technologie zu reduzieren, wäre eine Beschränkung auf möglichst wenige linguistische Beschreibungsebenen, wie es z.B. bei Rechtschreibprogrammen möglich ist (s.a. *Thelen* 2002).

Grundsätzlich lassen sich in diesem Bereich Techniken wie Präsuppositionsverarbeitung, Partnermodellierung und Generierung (vgl. Unterkapitel 3.6 und 3.7) nutzbar machen.

Während die beiden erstgenannten Systeme auf Bildelemente verzichten, werden diese in PROMISE gezielt eingesetzt, um den Unfallhergang per Bildabfolge zu spezifizieren und damit die faktischen Gegebenheiten präzise abzustecken und um die Unfallprotokollierung des Polizeibeamten im Dialog mit dem Lerner realistischer zu gestalten und per Unfallskizze die Bedeutung der Lernereingaben kenntlich zu machen. Dies ist besonders wichtig, weil die Erstellung der Unfallskizze dem Gebrauch räumlicher Präpositionen gewidmet ist.

Ein Großteil der Erfahrungen aus PROMISE ist in das in *Krüger und Hamilton* 1997 vorgestellte RECALL-System (Repairing Errors in Computer Assisted Language Learning) eingeflossen, welches intelligente und individualisierte Fehlerrückmeldungen in einem adventure-artigen Rollenspiel zum Schwerpunkt hat.

**Toolartige Systeme:** Am LATL (Laboratoire d'Analyse et de Technologie du Langage) werden ausgereifte, vor Ort entwickelte NLP-Technologien adaptiert, um sie in Lernwerkzeugen nutzbar zu machen (*Vandeventer und Hamel* 2000). Das syntaktische Generierungssystem GBGen, das auf der Theorie von Government and Binding aufsetzt (vgl. Unterkapitel 3.4), wird genutzt, um mit fortgeschrittenen Französischlernern in einem strukturierten Dialog Sätze zu produzieren und auf diese Weise die dazu erforderlichen Entscheidungen bewusst zu machen. In dem Prototypen SAFRAN (Système pour l'Apprentissage du FRANçais) werden – aufbauend auf dem syntaktischen Parser FIPS und dem Anzeigesystem FIPSGram – erwachsenen Französischlernern mittleren bis fortgeschrittenen Niveaus Satzstrukturen angezeigt. Das Sprachsynthesemodul FIPSvox (vgl. Unterkapitel 5.7), das ebenfalls auf den Analyseergebnissen von FIPS aufsetzt, fungiert als Aussprachereferenz und kann vom Lerner als Experimentierumgebung und zu Selbstevaluationszwecken eingesetzt werden. Darüber hinaus wird es für die phonetische Robustheitssteigerung des Lexikonzugriffs nutzbar gemacht.

Im Bereich elektronischer Lesehilfen stellen COMPASS (*Breidt und Feldweg* 1997), DEFI (*Michiels* 2001) und Glosser-Rug (*Nerbonne et al.* 1998) erfolgreiche ICALL-Tools dar. Das Ziel aller drei Systeme besteht im Grunde darin, aus elektronisch vorliegenden fremdsprachlichen Texten heraus per morphologischer Analyse (vgl. Unterkapitel 3.2) intelligente Zugriffe auf Lexikoneinträge zu bewerkstelligen. Den ersten beiden Systemen ist gemeinsam, dass sie komplexen lexikalischen Ausdrücken besondere Beachtung schenken und versuchen, über die Wortart hinausgehende Disambiguierungen vorzunehmen. Die besondere Leistung von Glosser-Rug besteht darin, dass es zu einem selektierten Wort zusätzlich zum Wörterbucheintrag auch (bilinguale) Korpusbelege anzeigt. Das System LogoTax (*Ludewig* 2001) teilt mit COMPASS und DEFI den Gegenstandsreich der Kollokationen als einen Spezialfall von Mehrwortausdrücken. Wie bei Glosser-Rug spielen allerdings auch bei LogoTax Korpusbelege eine zentrale Rolle (vgl. Unterkapitel 4.1). Alle vier Systeme beinhalten NLP-Komponenten, die in anderen Kontexten entwickelt wurden. Während die drei anderen Systeme

jedoch die Textrezeption fördern und dazu auf vorhandene lexikalische Ressourcen zugreifen, unterstützt LogoTax den Lerner, indem es den korpusbasierten Aufbau eines auf seine individuellen sprachproduktiven Bedürfnisse zugeschnittenen Lexikons für komplexe Wortverbindungen ermöglicht, denn im Gegensatz zu Idiomen bereiten Kollokationen dem Fremdsprachenlerner nicht so sehr beim Sprachverstehen Probleme, als vielmehr beim Produzieren sprachlicher Äußerungen.

Die aktuelle Version von LogoTax beschränkt sich auf V-N-Kollokationen wie *eine Vorlesung halten* oder *etwas in Frage stellen*. LogoTax unterscheidet sich dadurch von aktuellen Konkordanzprogrammen im CALL-Bereich, dass durch die Verwendung eines Parsers mehr syntaktische Kriterien in den Filterprozess einbezogen werden können. Dies steigert die Qualität der identifizierten Belege erheblich und ermöglicht es, einen Eintrag zu erzeugen, der in seiner Detailliertheit die Einträge herkömmlicher Wörterbücher maßgeblich übertrifft, indem er diverse Merkmale der V-N-Kollokation beschreibt und zu einem linguistischen Merkmal die entsprechenden Belege angezeigt werden können (z.B. alle Passivverwendungen). Auf diese Weise wird die Kluft zwischen generalisierendem Lexikoneintrag und konkreter textueller Verwendung überbrückt.

### 5.10.3 Fazit

Als Forschungsprototypen existieren sowohl isolierte ICALL-Bausteine als auch umfassendere Entwürfe für ganze Sprachlehrszenarien. Ferner ist zwischen tutoriellen und toolartigen Systemen zu unterscheiden. Die Prototypen zeigen in der Regel überzeugend, welche beachtlichen Verbesserungen unter Nutzung von NLP-Technologie im ICALL-Bereich denkbar sind. Dass es bisher dennoch kaum marktreife ICALL-Systeme gibt, hat insbesondere folgende Gründe: Es gibt wenig Erfahrungen hinsichtlich der Weiterentwicklung experimenteller ICALL-Systeme zu real einsetzbaren Systemen. Erschwerend kommt hinzu, dass viele ICALL-Systeme Neuimplementationen darstellen, d.h. nicht auf bereits zuvor existierende ausgereifte NLP-Module zurückgreifen. Ferner werden die didaktisch und somit anwendungsinhärent motivierbaren Einschränkungsmöglichkeiten des behandelten Sprachfragments häufig noch nicht optimal genutzt. Was die produzierten Analysefehler betrifft, ist anzumerken, dass sie vermutlich in einem Tutorsystem schwerer zu akzeptieren sind als bei einem Lerntool. Einem Tutor hängt indirekt der Anspruch an, die Lerninhalte umfassend zu beherrschen. In einem konstruktiven Lernszenario wird dagegen indirekt ein offener Lerngegenstand unterstellt, so dass es nicht wirklich verwundert, wenn ein Tool fehlerhafte Ergebnisse produziert. Darüber hinaus erfordert der Übergang von Forschungsprototypen zu real einsatzfähigen ICALL-Systemen eine intensive Zusammenarbeit mit Sprachdidaktikern, um den didaktischen Nutzen der NLP-Technologie voll ausschöpfen zu können. Und der hohe Entwicklungsaufwand schmälert die Wettbewerbschancen in einer marktwirtschaftlich angespannten Situation, in der Verlage nur sehr kurzfristige Planungen anstellen können.