36. NATURAL LANGUAGE PROCESSING IN COMPUTER-ASSISTED LANGUAGE LEARNING

John Nerbonne

Abstract

This chapter examines the application of natural language processing to computer-assisted language learning including the history of work in this field over the last thirty-five years but with a focus on current developments and opportunities.

36.1 Introduction

This chapter focuses on applications of computational linguistics (CL) techniques to **computer-assisted language learning** (CALL). This always refers to programs designed to help people learn foreign languages, e.g., programs to help German high-school students learn French. CALL is a large field—much larger than computational linguistics—which no one can describe completely in a brief chapter. The focus of this chapter is therefore much narrower, viz., just on those areas of CALL to which CL has been applied or might be. CL programs process natural languages such as English and Spanish, and the techniques are therefore often referred to as natural language processing (NLP).

To preview this chapter's contents, we note that NLP has been enlisted in several ways in CALL, e.g., to carry out the **alignment** of **bilingual texts** (see Chap. 24 on corpora) so that a learner who encounters an unknown word in French can seen how it was rendered in translation; to provide **lemmatized** access to corpora for advanced learners seeking subtleties unavailable in grammars and dictionaries (see Chap. 2 on morphology for an explanation of lemmatizing); to provide morphological analysis and subsequent dictionary access for words unknown to readers (Chap.2, Chap.18); and to parse user input and diagnose morphological and syntactic errors (Chap. 4,Chap. 12). Speech recognition has been used extensively in pronunciation training, and speech synthesis for single words (Chap. 16-17). The work on speech in CALL falls outside the scope of this

chapter, but it is noted since we expect its use in CALL to enable further deployment of NLP, perhaps in parsing spoken input.

The chapter deliberately ignores two further areas which arguably involve language learning, first, software designed to diagnose and assist the verbally handicapped (Miller and Klee 1995, Perry 1995, Connolly 1997, Somers 1999), and second, programs for assistance in developing and polishing advanced writing skills. The first area seems genuinely extraneous (there is little overlap with CALL), and the second has attracted little interest from NLP practitioners even though it enjoys substantial attention in the larger CALL literature (Pennington 1999, Schultz 2000).

36.2 Motivation

The traditional methods of learning and teaching foreign languages require a teacher to work 60 to 100 hours to bring an adult to a level at which she can function minimally with the foreign language (FSI 1973). By 'minimally functioning' we just mean being able to use short, pre-packaged phrases to communicate simple thoughts with long pauses and very foreign pronunciation. In general progress beyond this level requires a great deal more time and energy, roughly doubling at each proficiency level (FSI 1973). Very high levels of fluency are normally unrealistic goals. Naturally, some learners *do* reach high levels of fluency, even near-native levels, but they have to arrange for extensive contact beyond language instruction.

CALL will be an attractive alternative whenever these teachers are unavailable (within reasonable travel), inconvenient (due to scheduling difficulties), or unaffordable. CALL can also be a useful supplement in all of these situations (i.e., when more extensive direct instruction might be contrasted). Looking ahead to the role we will suggest is most promising for NLP in CALL, we note that even extensive experience with a foreign language is of little didactic value if it is incomprehensible to the language learner. Unguided "immersion" into a situation dominated by a foreign tongue does not result in proficiency. For learning to occur, learners must understand the language they are exposed to. A central role for CALL, especially when no teacher is available, is to provide comprehensible foreign language material in spoken and written form; to help

students understand foreign-language material, e.g., to verify that inflections carry particular grammatical import; and to provide exercise and test material. Because the computer can process information, it can provide such verifying information much more extensively (and conveniently) than any static source could hope to. The section below on **lemmatization** (36.6.4) illustrates this in the greatest depth.

A standard question on CALL which is reflected in recurring comments to net lists (see the end of this chapter for a pointer to one of these) asks whether CALL has been proven superior to traditional methods of language learning and learning teaching. Naturally, there have been studies aimed at measuring differences, but it is important to note that this challenge misses the main features of CALL which make it an attractive alternative or supplement to traditional methods of language learning and learning teaching. Even if it were *worse* than traditional methods, CALL would be used extensively—human teachers are not available in the numbers and budget ranges needed.

36.2.1 Inflated Expectations

This is emphatically not to say that CALL will ever provide all that its most wide-eyed admirers expect. Many potential primary users of NLP-enhanced CALL (learners) as well as secondary users (teachers and administrators) are unsophisticated in their expectations (Atwell 1998). Atwell (1999), while advocating increased user involvement in software development for English Language Teaching (ELT), illustrates how much deflation may be necessary:

> [...] it turned out that one of the main things that English language teachers asked for was a 'conversation practice machine': a computer system with which a learner could 'chat' in ordinary English, to practice their everyday conversational skills. [...] Another thing which ELT users seemed to want was an ability to correct errors and give guidance on the cause of errors in learners' conversations. (Atwell 1999, pp.31–32)

It looks as if the simple Turing test is not sufficient: we need a machine which not only converses naturally, but which spots errors and diagnoses their causes for less fluent

interlocutors! Teachers with some experience of CALL are naturally more sober. Levy (1997, Chap. 5) reports on an extensive survey conducted under CALL practitioners at universities. As a whole, this group is guided most by availability and reliability of software (Levy 1997, pp.148-150).

36.2.2 Application Sectors

Although we normally think of language education as a task for schools and universities, which it certainly is, there is an enormous need for language training outside of the schools. It is convenient to consider three distinct sorts of language learners and learning situations, as they differ in their interest for technical improvement and in their readiness to experiment, which in turn impacts on the assessment of application opportunities for CALL:

academic This group includes the schools and universities. Even though this is the group one most immediately thinks in connection with language teaching and learning, it has well-known, peculiar difficulties, e.g., penurious budgets, uncertain hardware and software infrastructures, and teachers who feel pressed to meet regular responsibilities without assuming the extra challenges of trying new technologies.

industrial Government and private industry organize their own language courses, often at considerable expense. We group specialized language schools under this category as well, but with the reservation that their means are often more limited, since their pupils often pay the bills directly.

self-study There is a sizeable number of people who study languages without the benefit of formal instruction organized by their jobs or educational institutions. They have stimulated a surprisingly large market in what is sometimes disparagingly dubbed "edutainment", but which proves that there is a mass market for CALL, as well.

Fox (1996, p.9) estimated the multimedia market in Europe at 1.9 billion Euro in 1994, 28% of which was educational. The CALL market is estimated as 20% of the latter, or about 106 Mil. Euro in 1994. The US market is more than twice the size

of Europe's with a similar proportion devoted to education and language. A threefold growth by 1999 was predicted. This leads one to predict a one-billion Euro business area by 2000. Meanwhile the CALL business benefits like all of Information and Communication Technology (ICT) from steeply falling price/performance figures, and from large-scale commercial interest in the underlying technologies.

36.3 Second Language Learning

In this section we turn to what is known about how people learn foreign languages, and how computers have been enlisted to help in this task, focusing first on how CALL *sans* NLP has progressed.

Applied linguistics studies **second language learning** and **foreign language learning**. Although the terms are used by some interchangeably, a distinction is often drawn. Foreign language learning normally takes place in classrooms and always takes place remote from extensive natural opportunity to use the foreign language. Second language learning occurs in a natural environment, normally in a country where the language is spoken. 'Second language learning' is often used as a cover term for the two situations, and that is the intention of this section's heading.

There are researchers who prefer the term second language *acquisition*, because 'acquisition' (as opposed to 'learning') emphasizes the degree to which automatic processes may play a role in the more natural situation when a language from the immediate environment is adopted.

Second and foreign language learning share an applied focus as scholarly fields: both consistently research not only how language learning normally proceeds, but also how it succeeds best. They seek to optimize learning, naturally with respect to the goals of language learners (e.g., scientific literature, tourism, or commerce), their (linguistic) backgrounds, and their age and educational level. van Els, Extra, van Os and Bongaerts (1977) is an excellent reference on issues in this branch of applied linguistics, and Dulay, Burt and Krashen (1982) a more theoretically oriented text on what research on second language learning acquisition reveals about underlying linguistic abilities and learning processing. Other texts can also be valuable (Larsen-Freeman and Long 1991, Ellis 1994).

36.3.1 The Fundamental Pedagogical Principle

A principle on which the different schools agree is that the material to which learners are exposed must be comprehensible to the learners in order for learning to proceed optimally ((Krashen 1982), cited often, *inter alia* by Widdowson (1990, p.111)). Krashen and others have gone on to postulate that the comprehensible input is the primary determinant in second language acquisition, dubbing this the "fundamental pedagogical principle" in second language teaching. Krashen's idea is that comprehensible input triggers automatic processes which are ultimately responsible for successful language learning. There is an important qualification and an important restriction in scope. Krashen (1982, p.66) explicitly qualifies his postulate that the comprehensible input determines second language acquisition, noting that "conscious learning" (i.e., morphology and "easy" rules [scare quotes in original–JN]) does contribute to second language proficiency. The restriction in scope is implicit, but there is no suggestion that exposure to comprehensible input is sufficient for learning writing, even though many second-language learners do aspire to competence in writing.

Krashen's writings are clearly argued. He justifies the overarching postulate by showing how it makes sense first of the fact that understanding always precedes production—this is so because understanding is a prerequisite for production. The postulate also helps explain the variable speed of children vs. adults as language learners. Adults comprehend more initially because they make better use of non-linguistics cues, etc., so they learn more quickly than children initially. The fact that children are ultimately more proficient is attributed to the adults' "filtering" foreign input for emotional reasons. He reviews a series of studies on instruction and is able to generalize that the amount of comprehensible input was a determining, often confounding variable in these. Carter ([1]1987, 1998, p.208) makes the point that lots of vocabulary learning must take place without explicit teaching—since it isn't taught. Krashen develops consequences of his ideas for the importance of reading (Krashen 1989), and the use of media and language laboratories, and the combination of language teaching with specialized subject matter (in some circumstances). Krashen has modified his original views to admit that a special "weak interface" brokers the learning of second languages, and that this distinguishes it

from learning first languages (Krashen 1985).

There are nonetheless important recent reactions to Krashen's emphases on automatic processes of acquisition. DeKeyser (1995) has shown that adult learners very definitely benefit from explicit instruction in grammar rules, claiming more strongly that adult learners basically do not learn to speak grammatically without explicitly being taught the rules. DeKeyser (2000) shows that the very few adults who appear to learn successfully without explicit instruction score high in analytical ability (which is irrelevant in children's success in foreign language acquisition). This suggests that they are exceptions precisely because they are capable of abstracting the rules themselves, to some extent. Norris and Ortega (2000) review dozens of studies over the last thirty years, and claim that Krashen had it exactly wrong, at least when it comes to adults' learning of grammar. Naturally, the range of potential CALL applications is broader if adults need explicit grammar instruction.

The applied linguistics community agrees further to a great degree on fundamentals such as holding the attention of learners, encouraging repetition, and aiming for varied practical exercise. There is less agreement on issues such as the value of formal grammatical tutoring (as just noted), the value of correcting errors, the time at which to encourage speaking, etc. See (Larsen-Freeman and Long 1991) for more on the range of (conflicting) ideas current in language pedagogy. Under the circumstances it is wise for CALL developers not to embrace any theory too exclusively, but to remain consonant with different approaches (cf. Lantolf (1996)).

It would seem straightforward that all that is known about second-language learning ought to inform the development of CALL materials. It provides a modest amount of theory, and it suggests useful concepts and categories (natural learners vs. taught learners, reading/listening vs writing/speaking), tests and means of evaluation, simple techniques, interdependencies between media, techniques, curricula, etc. But surprisingly, one respected recent book on CALL (Levy 1997) takes pains to consider alternative theoretical orientations for CALL. Levy's primary goal is to show that CALL should be regarded as a "tool", in language learning rather than a "tutor", or substitute for a teacher, and in general CALL practitioners have been receptive (Wolff 1999). But on the way to this

goal, Levy considers Psychology, Linguistics, Artificial Intelligence and other fields as candidates within which to locate CALL in an extended diatribe meant to establish that poor "conceptualization" is hampering progress in the field.

There are no grounds to think that second-language learning is fundamentally different when the computer is involved, however, and this suggests that CALL be viewed simply a technical innovation in second-language learning. And Levy (1997, Chap. 8) finally also comes to the sober—and, I believe, correct—view of CALL as an applied field that ought to pay more attention to the concrete success of its products and less to its theory and its ultimate place in the scientific scheme of things,[1] so perhaps the extended discussions in early chapters are intended to convince the unreflective.

36.4 CALL

CALL seeks to employ computers in order to improve language learning. CALL spans the range of activities in language pedagogy — hearing, speaking, reading, and writing — and draws from nearly all areas of ICT. Even if most CALL applications are automated language exercises, exploiting hypertext, simple database and network technology, and digital audio and video, one finds many others, including ingenious applications of everyday technology such as word-processing and email. Levy (1997) is valuable for its survey of the surprisingly long history of CALL, reports on the field's extensive reflection on its proper relation to applied linguistics, computer science, and psychology, and his own astute view of its proper, technology-driven nature in the final chapters. There is no mention of opportunities for several common CL techniques, however, e.g., **alignment** software or software for **lemmatization**.

Although CALL employs the computer to assist in language teaching and in language self-study, it primarily uses non-language technology, as noted above. The basic ICT technologies used are as follows:

(simple) database technology to record and present student work; and

[1] On this point, Levy's view may be contrasted with Holland's (1995, p.x), who sees "theory shaping technology." Levy argues that effective CALL programs have seldom been developed with much direct inspiration from theory and that more practical inspiration has been crucial.

digital audio and video to vivify examples of language use (Rothenberg 1998);

hypertext to provide varied access to exercises and explanatory material (Jager 1998);

network communication to bring learners and teachers into easier contact (Hoffman 1996, Holliday 1998). An ingenious use of email is to put intermediate learners in symmetric contact — native English speakers learning Spanish with native Spanish speakers learning English (Appel and Mullen 2000*b*, Appel and Mullen 2000*a*).

We note this work to give an idea of the sorts of efforts which compete for development resources with NLP-informed CALL. As the introduction warned, this chapter will not discuss the non-NLP work. Cameron (1999*b*) is a good recent collection of articles on CALL which give a sense of the range of approaches that are used, current issues, evaluation techniques, etc. Several of its articles discuss CL techniques directly, however, and these figure below. Levy (1997) is a more reflective work, and indispensable for some of its comments on the history of the field.

36.4.1 History of CALL

CALL looks back on a long history of serious and widely interdisciplinary work. Levy (1997, Chap.2) reports *inter alia* on PLATO ('Programmed Logic for Automatic Teaching Operations'), a University of Illinois project begun in 1960, and TICCIT ('Time-Shared Interactive Computer-Controlled Television'), a Brigham Young University effort started in 1971. Both were sophisticated, featuring 'talk' facilities for exchanging typed messages; exercises for reading, writing and listening; opportunities for instructors to modify material; and self-paced organization of materials. Evaluation of this early work noted that while the performance of students was good, the student retention rate was not. The systems were not popular with students, and this remains a caution to current work (Hamilton 1998). The 1983 Athena Language Project (ALP) started at MIT in 1983 and involved artificial intelligence tutoring techniques and multimedia to provide instruction in five European languages. In comparison to the earlier projects ALP focused more on communication as opposed to grammar, and it included games and puzzles (Murray 1995, Felshin 1995).

There is a new web site at `www.history-of-call.org`with some presentational glitches (in Netscape as of 3/2001) which contains many very useful pointers for those interested in the history of this technology.

36.5 Natural Language Processing

As earlier chapters have noted, NLP focuses on how computers can best process language—analyze, store, sort and search it. It seems natural that NLP should be applied to the task of helping people learn language.

36.5.1 NLP from the CALL Perspective

NLP offers a range of techniques for processing natural language; these are explored in earlier chapters. A few successful CALL programs make use of this language technology, but the great bulk do not (Zock 1996). Since CALL without NLP technology is successful, we must ask how (and whether) language technology can be applied to improve language learning and language teaching.

There are voices which urge caution in applying NLP to the CALL problem. (Salaberry 1996), p.12 assesses the suitability of language technology for CALL quite negatively, but not atypically:

> Linguistics has not been able to encode the complexity of natural language [...] That problem has been acknowledged by the most adamant proponents of Intelligent CALL [ICALL (JN)]. Holland (1995) lists the reasons that have prevented ICALL from becoming an alternative to CALL. The most important reason for this failure is that NLP (Natural Language Processing) programs—which underlie the development of ICALL—cannot account for the full complexity of natural human languages.[2]

Salaberry attributes the unpopularity of NLP methods in CALL to NLP's failure to produce full-fledged intelligent dialogue in unrestricted form. The earlier chapters of this book should make it clear many NLP techniques are extremely reliable and suitable

[2] Salaberry's reference to Holland (1995) is not accompanied by bibliographic information.

for demanding, robust application (see Nerbonne, Jager and van Essen (1998) for more extensive discussion of this negative perception of NLP among CALL developers, and see Salaberry (2000) for a thoughtful rejoinder). We note here only that the negative perception exists, and that, given that the great bulk of existing CALL programs and packages make little use of language technology (Zock 1996), one should not simply regard CALL programs by definition as applications of NLP.

To end this section on a more positive note, some language learning experts note the need for NLP techniques in CALL very clearly, if implicitly. John R. Allen developed CALL programs for over twenty-five years, and Allen (1996/1997) presents his view of how CALL must develop, focusing on the issue of how we can fashion useful reaction to student input. Among other needs, he discusses clearly how simple regular expressions will fail as specifications of syntactic patterns (p.446), and he calls for better specification techniques of the sort that would be bread and butter to all NLPers. He further notes that CALL needs better techniques for recognizing parts of speech and morphological variants (p.452), again calling for techniques which NLP can normally provided reliably. See Section 36.6.4 for a discussion of NLP work in CALL which uses both part of speech tagging and lemmatization software. Allen is likewise clear about the need to incorporate intelligent tutoring schemes of the sort which Henry Hamburger and Kathleen McCoy have investigated most extensively (see Section 36.6.6 for more on this work).

36.5.2 NLP Interest in CALL

There has been regular and growing interest in the NLP community in the CALL application area.

Appelo and de Jong (1994) sparked the interest of a number of NLP researchers in CALL. Holland and Kaplan (1995) surveyed the state of the art in 1995, and Holland, Kaplan and Sams (1995) surveys lots of work, which focuses on the question of how to use NLP in a pedagogically responsible way, and on using tutoring techniques from artificial intelligence (see 36.6.6 below, as well). More recently, an Applied Natural Language Processing conference devoted a session to CALL (*Proceedings of the Fifth Conference on Applied Natural Language Processing* 1997). The session was organized

by Melissa Holland and Henry Hamburger, and it featured papers by researchers in NLP (Schoelles and Hamburger 1997, Nerbonne, Karttunen, Paskaleva, Prószéky and Roosmaa 1997, Dorr 1997). In that same year Jager, Nerbonne and van Essen (1998) organized a conference devoted to 'Language Teaching and Language Technology' whose proceedings appeared in 1998. They explore especially the opportunities for language technology in CALL, and include several reports on CALL applications that exploit NLP, including **alignment** techniques (Paskaleva and Mihov 1998), **lemmatization** (Dokter and Nerbonne 1998, Roosmaa and Proszeky 1998) and **parsing** (van Heuven 1998, Murphy, Krüger and Grieszl 1998). Schulze, Hamel and Thompson (1999) is the collection of papers from a UMIST conference that was organized in 1998 (Schulze and Hamel 1998). Olsen (1999) presents the proceedings of a joint workshop sponsored by the Association for Computational Linguistics and the International Association for Language Learning Technology. If the booming interest in the speech community is a good indication, we should see further growth in the NLP community's interest, too.[3]

36.6 Contributing NLP Technologies

This section presents work in NLP which has sought to contribute to CALL. Many technologies have been enlisted in an effort to make illustrate linguistic structures, make language comprehensible, provide varied exercise material, and spot and correct errors.

concordancing Concordance programs are by now so well known and widely available that one might hesitate to include them under NLP, but they are language processing programs, and they have inspired an enthusiastic group of advocates among language teaching professionals.

text alignment When alignment programs are applied to align bilingual texts, the resulting parallel texts offer a wealth of information to advanced language learners. There has been a stable market for bilingual texts in language instruction for decades. See Chap. 24.

[3] Bernstein, Jong, Pison and Townshend (2000) and other papers in that collection provide a good overview of work on speech in CALL.

speech recognition and synthesis Speech technology is applied to generate pronunciations, particularly pronunciations of isolated words. This relieves language learners of needing to understand pronunciation transcription systems. Speech recognition also is applied to check (and improve) pronunciation. See Chap. 16-17.

morphological processing Lemmatization, but also morphological generation has been deployed to provide drill material for learners, to facilitate dictionary lookup of words (which is impossible in inflected languages without morphological analysis), and to make corpus access more flexible. See Chap. 2 and Chap. 18.

syntactic processing Syntactic generation may be employed to create exercise material, and parsing is employed both in order to clarify linguistic structure, but also in order to spot and diagnose errors in learners' output. See Chap. 4 and Chap. 12.

Machine translation (MT, see Chap. 27) has been argued to involve most of the language technology needed for CALL,(Levin 1993) except perhaps error spotting and diagnosis, but we have noted the underlying techniques here rather that the analogy in application. MT indeed offers the application builder in CALL a great deal, but that is also because it is a very complex application.

The following sections focus, not on technologies, but rather on the sorts of tasks to which they are put in CALL.

36.6.1 Pronunciation

Speech technology is treated in Chap. 17-18, but as it is not the focus of this book, we recall here merely that the successful use of speech technology in CALL will generate more interest in NLP. In fact, there have been commercial products using speech technology in CALL for some time. Martin Rothenberg, known to phoneticians as the developer of the "Rothenberg mask", a device for capturing air pressure information sensitively, is also the founder of Syracuse Language Systems, the world market leader in CALL products. If Rothenberg was early, he was not alone for long. DynEd and The Learning Company also offer pronunciation correction packages commercially (Aist 1999). Bernstein et al. (2000) measures the ability of a system marketed by Ordinate to score non-native pro-

nounciations for quality, noting a high overall correlation with expert human graders ($r = 0.94$).

Speech is seeing an even stronger growth in interest in CALL than NLP. The forty-five contributed papers in Carlson, Dunger, Granstrom and Oster (1998) were devoted to speech technology applied to CALL, and EUROCALL 2000 included a special workshop on "Intergrating Speech Technology in Language Learning", the fifth in a series. speech recognition as one of its focus areas (see the web site listed below in "Resources"). Holland (1999) reviews speech recognition in a range of CALL applications. Progress in this area will undoubtedly enable more NLP applications.

36.6.2 Corpora and CALL

In this section and the following, we examine the use of corpora in CALL (insofar as NLP is involved). Corpus-based CALL involves have students look into corpora in order to appreciate linguistic patterns and distinctions more elaborately, and it is argued to be of value to very advanced students. Language technology is already seen to contribute to language teaching as it uses corpora: Concordancing and lemmatization have been applied to corpora in order to try to facilitate the use corpora in CALL, and their added value is substantial.

There has been focused interest in using language corpora for CALL (Wichmann, Fligelstone, McEnery and Knowles 1997). Corpora are valued for providing access to authentic language use, unmediated by grammarians' theories, prescriptivists' tastes, pedagogy's traditions, or even lexicographers' limitations. It is very clear that corpora can only be useful for advanced students—beginners would understand simply nothing they saw.

There are moderate and extreme views on how corpora should best be utilized. The moderate view espouses the value of corpora, especially when accompanied by good search tools, for instructors and very advanced students—those for whom unabridged dictionaries and comprehensive grammars are insufficient as sources of information on nuances of meaning, common usage, or stylistic level. A more extreme view is often attributed to what Tim Johns has dubbed 'data-driven learning', which emphasizes the

role of discovery in the language classroom, facilitated by tools for corpus analysis. Johns and King (1991, p.2) finds that "the language learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data".

Of course, a bare corpus is of reduced value to both teachers and students. If a corpus is to be used effectively, then it ought to be supplemented with at least two tools from language technology:

concordancing Concordance programs of the sort that finds words in texts and presents them perspicuously in context have inspired an enthusiastic group of advocates among language teaching professionals (Wichmann et al. 1997).

... pointing an accusing	**finger**	at governmental agencies ...
... who had injured one	**finger**	and sustained bruises ...
... arrested for giving the	**finger**	to an unsympathetic ...

lemmatization In inflected languages, such as French, Spanish or Russian, the chance of finding any inflected form of a given lemma (i.e., searching for any form whatsoever) is two orders of magnitude greater than the chance of finding a particular inflected form. The chance can be even worse if the form is unfortunately chosen.

These tools should reduce the amount of preparation which instructors may need in order to make use of corpora effectively, a problem which Holmes (1999, pp.253-256) warns about. Holmes suggests that the apparent successes of corpora in CALL may be due to the enthusiasm of the innovators rather than the intrinsics of the activity or technique. He noticed no improvement in students' ability after their experiments with corpora (p.256). This is not encouraging, but it should be clear that language learners need to invest a good deal of effort in corpora in order to obtain the information that might have been distilled into a brief dictionary entry or standard grammar explanation.

36.6.3 Bilingual Corpora

The fundamental reason to explore bilingual texts in CALL is that they grant the language learner the same access to authentic language use, only now accompanied by convenient translation into a known language. This increases the chances, of course,

that the foreign-language corpus material will be comprehensible to learners, which, as noted above, is one of the prime requirements of all effective foreign-language pedagogical material (Krashen 1982). The advantages of immediate access to genuine material thus accrue to language learners with access to bilingual texts, but now with the added advantage of comparison to a language they are assumed to know. Barlow (1996) illustrates these advantages by displaying the results of searches for English reflexives, on the one hand, and the English lexeme 'head', on the other. His examples show that French reflexive patterns mirror English only partially, sometimes using reflexive pronouns (allowed himself *s'est laisse*), but often omitting them (buy themselves lunch *acheter un déjeuner*), or using an impersonal construction (enjoyed himself *l'enchanta*), or, in some cases, using wholly different lexical material (speaking for myself *en mon nom*). The reflexive example is particularly striking in light of the extensive grammatical analyses that have been devoted to reflexive pronouns. Barlow's example suggests either that the rules put forward by such analyses fall short of providing adequate guidelines for language learners seeking full mastery of the language, or that the role of lexis is more extensive than often supposed. The example of Eng. 'head' is of a sort more familiar to language learners: it is easy to find several common French equivalents, including *tête, chef*, and *directeur*, as well as to show that idiomatic uses show up frequently ('head on', 'keep one's head down').

Translation is often a course of study for advanced language learners, and Peters, Picchi and Biagini (1996) note that as the goal of translation has shifted from formal linguistic equivalence to pragmatic equivalence, the bilingual corpus has risen in importance vis-à-vis the bilingual dictionary. The dictionary can never vie with extensive corpora in cataloguing and illustrating the sorts of correspondences found in translation. Danielsson and Ridings (1966a) report on an educational tool used in a training program for translators. It is based on Danielsson and Ridings's (1966) parallel corpora work, and it is based on one million words which are aligned at the sentence level. Students of translation benefit from the abundance of material which they use to find unusual translation equivalences. To summarize this section: a number of researchers have begun experimenting with bilingual corpora in language learning situations, and they advocate

more extensive experimentation. They adduce convincing reasons why bilingual corpora supply information that would otherwise be unavailable. They note unanimously that the use of bilingual corpora only makes sense if good software is available to support the sorts of searches which instructors and students wish to conduct. At the same time, we must note that the field is very young. There is little report on actual uses of bilingual corpora by students, and the (extensive) reports by instructors may be of interest more for their contributions to comparative grammar and descriptive linguistics than for their contributions to language pedagogy. There have been few attempts to evaluate the effect of the use of parallel, bilingual texts on language learning.

Bilingual corpora enjoy in some sense a special status since manually aligned and manually prepared texts have been used for decades in language teaching (even if they are not the object of current research interest in language teaching). Parallel texts show translation near originals, and they are a reasonable guarantee that textual material will be comprehensible, in accordance with Krashen's dictum, noted above. Of course, NLP can provide technology to support the use of bilingual texts.

text alignment Unaligned bilingual corpora offer little practical value to teachers and students. Once they are large enough for there to be a good chance of finding particular words or constructions, they are too large for simple search mechanisms to be effective (Nerbonne 2000)

Corpora are undeniably valuable to advanced learners, but in a user study with second-year students who were given automatic access to an online dictionary, a description of the grammatical function of inflectional morphemes, and examples of words from corpora, students made little use of the aligned, bilingual corpora (Dokter, Nerbonne, Schurcks-Grozeva and Smit 1998). Of course, they were not advanced enough to need to go beyond the dictionary, and they had been instructed that they would be asked about the content of a reading passage, which arguably did not promote an "exploratory" attitude toward the software package. Dictionaries and brief grammars are efficient means for learners to acquire basic information about words and construction. They may be the language technology of the 17th century, but they have proven their worth.

36.6.4 Morphology

The idea of applying morphological analysis to aid learners or translators, although not new, has not been the subject of extensive experimentation. Goethals, Engels and Leenders (1990) report on A.D.A.M. & E.V.E., a software package for advanced learners which generates practice material from a corpus of authentic texts. The exercises stimulate practice in different morphological forms and are generated in a way sensitive to frequency. Antworth (1992) applied morphological analysis software to create glossed text. But the focus was on technical realization, and the application was the formatting of inter-linearly glossed texts for scholarly purposes. The example was Bloomfield's Tagalog texts. Kempen (1992) and Kempen and Dijkstra (1994) explore morphology as a means of recognizing and correcting spelling errors which seem to be triggered by morphological insensitivity. French has inflected forms which do not vary in pronunciation even though they are distinct in morphology and spelling, e.g., French *parler, parlais, parlait,* and *parlez* and Dutch final <t> and <d> are both pronounced [t] even though they differ in morphological significance.

The COMPASS project (Breidt and Feldweg 1997) focused on providing "COMPprehsion ASSistance" to less than fully competent foreign language readers. Their motivation seems to have stemmed less from the situation in which language learning is essential and more from situations in which one must cope with foreign language. In addition, they focused especially on the problems of multi-word lexemes, examples such as English *call up* which has a specific meaning 'to telephone' but whose parts need not occur adjacently in text, such as *call someone or other up.*

GLOSSER assumed learners would have the task of understanding texts and would call on GLOSSER for information on the words in the text (Nerbonne, Dokter and Smit 1998). Three sorts of information are provided, first a lemmatization of the inflected form, together with a brief indication of the grammatical significance of the morphology; second, the dictionary entry for the lemma; and third, examples of the word from corpora, some of which are bilingual. In order to improve the accuracy of the lemmatization, the text is labeled with part-of-speech (POS) tags (see Chap. 11) before it is subjected to morphological analysis (Bauer, Segond and Zaenen 1995). The

suitability of the NLP technology was assessed technically (Nerbonne and Dokter 1999) and from the perspective of users (Dokter and Nerbonne 1998). NLP improved the chances of students finding information in dictionaries (which is not always trivial in the case of morphological irregularity), and NLP improved the usefulness of corpora—lemma searches are much more effective than string searches, but it remains the case that GLOSSER's most valuable feature for users was that it made access to the dictionary fast (more than 100 times faster than learners' searches).

It is probably correct to view NLP's role in this sort of system more as a tool to be used when needed rather than as a supplying pedagogical content. GLOSSER supplies content about the grammatical meaning of morphology, but users were much more interested in the dictionaries. GLOSSER maintains an demonstration at `www.let.rug.nl/~glosser`(Nerbonne, Kleiweg, Smit, Kuipers and Dokter 2001). The demonstrator is reduced in order to protect proprietary modules.

Surely the main reason for the success of this sort of use of NLP is that morphological processing is sufficiently mature to support nearly error-free lemmatization (in a random sample of 500 words we evaluated, there were no mistakes in lemmatization, and 12 faulty POS assignments). This functionality can be used to automate dictionary access, to explain the grammatical meaning of morphology, and to provide further examples of the word in use (perhaps in different forms). Second language learners look up words faster and more accurately using systems built on morphological processing. It is to be expect that this will improve their acquisition of vocabulary.

The Ecology of NLP in CALL

It is convenient to include here some discussion of some wider issues vis-à-vis the design of CALL programs which utilize NLP. These issues involve both the immediate context of use for the programs as well as the larger question of how these need to fit into complete curricula. Naturally different kinds of uses of NLP will prompt different questions, but we want to clarify that the use of NLP, the user interface and the purpose of the CALL program are all of potential importance, and that there may well be research results in second language acquisition which can inform the particular deployment of NLP in

CALL.

GLOSSER embedded NLP in a natural task, namely reading. The language ability targeted for improvement was vocabulary acquisition. The idea was that students who could read more (because GLOSSER made the texts more accessible) would tap into Krashen's automatic processes of acquisition. This meant that one should see vocabulary increase not just, and perhaps not even primarily in the vocabulary that is looked up, but in the many words the learner has to process. The value of reading for vocabulary acquisition is widely supported by research in second-language acquisition. Carter ([1]1987, 1998, p.209) makes the stronger point that second language learners CANNOT learn sufficient vocabulary from vocabulary lists in textbooks—some sort of implicit acquisition must be taking place, and GLOSSER enables the learner to comprehend more easily than he otherwise would, providing the right sort of experience. Reading is seen to have a further advantage over vocabulary lists in that it introduces new vocabulary within a sensible context, even without the use of additional sources like dictionaries (Krantz 1991). Krantz (1991) emphasizes the importance of learning vocabulary words in context, and GLOSSER supports exactly that. The context provided by full text not only contributes to a better understanding of the possible uses of a specific word, but it also creates a framework in which words are more easily remembered (Mondria 1996).

A further perspective from second language acquisition is the extent to which programs exercise natural communication rather than mastery of grammar, and again from this *communicative* perspective it seems fair to rate GLOSSER as successful (Warschauer 1996). The choice of reading material is entirely up to the student and/or teacher, but it may include authentic materials, which Widdowson (1990) and others have argued improves the quality of learning by involving the learner more directly in the community in which the target language is spoken.

Hulstijn (1992) notes a further point of choice for second language learners who are acquiring vocabulary through reading, and that is whether they have immediate access to dictionary meanings or whether they try instead to infer meanings from texts (without using a dictionary). Hulstijn designed a series of experiments to measure vocabulary retention under these two conditions. With some qualifications, Hulstijn recommends

that dictionary meanings *not* be made available immediately, but rather that students be encouraged to try to infer the meaning from context first. This was a recommendation we finally did not implement in GLOSSER, conceding that learners needed to learn to use the tool intelligently. It could make an interesting follow-up study. Several further variants were investigated, e.g., one which allowed students to leave notes in the texts as they read.

The purpose of the discussion here is not to praise GLOSSER's stance as pedagogically sound, but rather to signal to potential developers (or users) of NLP in CALL that the immediate pedagogical context of the NLP machinery is as important as the technology in evaluating how well an application can succeed. A clever setting for NLP can make up for potential problems: so Sanders and Sanders (1995) embed their dialogue system in a game about uncovering a spy ring. The context justifies their systems refusal to deal with some words, which learners are told are "forbidden". Levin and Evans (1995) try to exploit sublanguages as a means of keeping processing within reasonable bounds of accuracy and efficiency. In their case the sublanguage was that of a first-year textbook that might have been used in the same course of instruction as their CALL system, and they present a lucid discussion of the options for systems whose linguistic coverage is less than perfect, finally opting for interactivity (to counter ambiguity) and some user preparation, the latter amounting to honest disclaimers regarding the possibility of errors.

This is also a convenient point at which to illustrate concretely another general point about CALL as an NLP application: the intended users (language learners) are intelligent and also undemanding. The users were in intelligent in using the information GLOSSER provides: even though the program did occasionally make errors, e.g., 2.4% of the POS categorizations were wrong, the users never stumbled over these. They are accustomed to using dictionaries and grammars, and they realize that the information they contain is not always relevant to the texts they have. The users were also undemanding: as experienced researchers in NLP realize immediately, GLOSSER suggests several further development paths, e.g. providing information about **multi-word lexemes** and attempting **sense disambiguation**. But since users are accutomed to dictionaries, that

is the normal comparison, and there were never questions about why those improvements were not available. This suggests that CALL could be a more useful vehicle for practical experimentation in applications of NLP than it has been to-date.

36.6.5 Grammar

Certainly most of the work in NLP-inspired CALL has been done on syntax. Michael Zock and his colleagues have developed systems which harness natural language **generation** (see Chap. 15) to walk language learners through exercises in sentence construction (Zock, Laroui and G.Francopoulo 1989, Zock 1992, Zock 1994). The learners are made aware of the linguistic rules which underlie sentence construction, and in particular about the effects that some choices in expression have, e.g., the effect that the grammatical number of a clitic pronoun has on the form of the French participle in the perfect tenses

Les filles! Qui est- ce qui les a aidé<u>es</u>?

The girls! Who has helped them?

If the plural clitic pronoun *les* had a masculine antecedent, then the participle would need to have the form *aidés*. If the clitic were feminine singular the appropriate form of the verbal cluster would be *l'a aidée*, and if the clitic were masculine singular, then the cluster would be *l'a aidé*. Zock's systems clarify such dependencies to langauge learners in an interactive fashion.

Work on generation has also been applied to suggest games with language learners (Pasero and Sabatier 1998), e.g., the game of constructing a sentence from a list of words, and also in order to generate dialogue, e.g., in the microworlds tutorials which Hamburger (1995) and others have pursued.[4]

Furthermore, a good deal of work has been using parsers and analysis trees to try to give students an idea of syntactic structure. Kempen (1999) presents a system for teaching students fundamental grammatical concepts as a *preparation* for foreign language study. The system is capable of providing different grammatical analyses—either dependency grammar style (see Chap. 4) or phrase-structure analyses (see Chap. 4), and

[4] See, however, Levin and Evans (1995, p.92) for a criticism of the microworlds approach. They question whether it will scale to useful size with a reasonable amount of work.

Kempen goes on to suggest appealing graphical renditions of these analyses—families of friendly "spooks" joining hands to indicate dependence, and multi-tiered "temples"" reflecting constituency.

Rypa and Feuerman (1995) try to support reading comprehension by providing analysis trees on demand. They focused on nontrivial grammatical points, and they use a Lexical-Functional Grammar (LFG, see Chap. 4) covering about 500 lexical entries, and a set of examples based on these. The most innovative aspect of their work involved the use of tree-based pattern matching in order to find examples of constructions satisfying a given description. Since they envisage that the system will be used by people naive about NLP, they needed to provide an interface to allow these users to specify what sorts of constructions they wish to find examples of.

Yet another use for syntactic analysis is suggested by Levin and Evans (1995), who present one of the most sophisticated systems exploiting NLP in CALL, ALICE-chan. Although it will be discussed here, it might also have figured in the section on morphology, and it performs error analysis, as well (discussed in the following section). Like Rypa and Feuerman's (1995) work, it is based on LFG, which it uses to parse examples for students, to create grammar exercises from examples and further specifications, and to parse user input. The presentation suggests that the system attained the full coverage it targeted, i.e., the **sublanguage** covered in its chosen first-year Japanese text. This sort of approach—taking care that NLP work dovetails with the overall language curriculum—undoubtedly bears emulation.

Finally, there have been other uses of parsers which are harder to classify, but certainly interesting. For example, Loritz (1995) uses a parser to classify the constructions used by deaf learners of English. The resulting analysis shows where errors are made, but also what constructions are avoided or overused by learners.

The topic of spotting and classifying errors in learner input is peculiar to CALL and has been the focus of so much attention that we devote the next, separate section to it.

36.6.6 Error Recognition and Diagnosis

A central technical issue in NLP applied to CALL has been that of recognizing, diagnosing and reacting to student errors. This certainly a natural concern—learners will always make mistakes. It is also a central theme in CALL quite apart from NLP. Allen (1996/1997) writes almost exclusively on error recognition in programs for language drill in a paper he entitles "Ten Desiderata for CALL"—as if the question of responding to errors was all that CALL consisted of. It is useful to mention the paper because Allen is a language teaching expert, and he is emphatic about the recommendation that learners should *not* be corrected on every minor error. It is useful to set aside as a pedagogical question just which errors should be pointed out and what sort of feedback is appropriate. Collaboration with language teaching experts is recommended on that issue, and we can focus here on recognizing and diagnosing errors.

If a student using a CALL program is asked to answer a question about a picture, and if in response to the question 'What do the children see?', she replies 'See dog.', then errors have been made, and the program will need to note that (even if not every error results in a reaction to the student). If there is to be intelligent feedback, then the program needs to hypothesize a source of error, in this case probably the failure to appreciate the English grammatical requirements that subjects need to be expressed, and that singular count nouns normally appear with an article. (The answer would have been grammatical if the student had replied 'They see a dog'.)

The first reaction of non-NLP experts on considering this problem is often to ask whether the error-spotting machinery found in grammar checkers might not be put to good use here. And CALL practitioners have experimented extensively with standard grammar checkers, of the sort developed in order to make recommendations to native speakers. Tschichold (1999) compares automatic grammar checkers systematically to the judgments of language teachers, and notes first that they are somewhat unreliable. Although they flag many more areas than expert humans, they often fail to spot problems, and what they flag as problems is a wide variety of errors, stylistic infelicities, and perfectly correct text. She concludes that that automatic grammar checkers have only a modest role to play for (advanced) students who understand their limitations (see also

citeasnountschichold:99a). Of course, some of the technology that has gone into grammar checkers could be interesting for CALL, e.g., the heuristics which Jensen, Heidorn, Miller and Ravin (1993) introduce to complete structures when straightforward parsing fails.

The initial reaction of many NLP experts is that error recognition and classification might still be too difficult for existing technology. First, parsing often turns up multiple analyses, and some of these could make sense of the erroneous input. If the parse has access to a large dictionary, then it will note in the example above (*see dog*) that *see* can be a noun (*the bishop's see*) and postulate a noun-noun compound analysis. In longer sentences one often encounters thousands of possible analyses. Second, even where one can determine that there is no available analysis, so that an error must be present, it goes beyond the bounds of standard NLP to try to classify the error as to its cause. And in general there will be multiple plausible hypotheses about the source of the error. Third and finally, some NLP experts are sceptical of what sounds like a user-modeling issue. Didn't Sparck Jones (1991) show that user modeling is still technically ineffective, even in much simpler settings?

However, there is still reason for optimism in the CALL case. First, although ambiguity rises steeply as a function of sentence length (see Chap. 12), very short sentences (of six words and fewer) tend not to be syntactically ambiguous. Second, as Michaud and McCoy (2000) note, it is possible to add rules to existing grammars which explicitly cover the errors users make. Schneider and McCoy (1998) attribute this technique to Weischedel, Voge and James (1978), and refer to it as "*mal*-rules". And third, if this is a user-modeling issue, it is an issue of a limited sort—the model only needs to follow the language learners' progress in acquiring the various linguistic rules and constructs (Michaud and McCoy 2000). McCoy, Pennington and Suri (1966) provide an extensive corpus analysis of errors made by deaf learners of English to support the claim that errors are (somewhat) predictable, which justifies the use of user model incorporating at least the native language of learners (and its error tendencies in the second language).

Constraint relaxation would appear to provide an alternative to the *mal*-rule technique of recognizing and diagnosing errors, but projects which have set out to apply it

seem to fall back on the use of special "error rules" (Murphy, Krüger and Grieszl 1998). Weinberg, Garman, Martin and Merlo (1995, p.43) claim to use a variant of constraint relaxation to diagnose errors, spotting errors by "modulating the exactness of the match between grammatical patterns and student input", but they do not provide technical details except to note (p.42) that the parser is tuned to seek out particular errors, e.g., subject-verb agreement. Logically, constraint relaxation would appear to be similar to the use of *mal*-rules, but technically it seems to suffer more from problems of effective processing (ambiguity).

Menzel and Schröder (1999) report on a system which uses both *mal*-rules and constraint relaxation. It considers all possible associations of words, and arrives at its best considered result by a process of elimination. Non-eliminated analyses are ranked according to the severity of the violation which the *mal*-rule encodes or, alternatively, according to the strictness of the constraint violated. In addition to standard grammar rules, Menzel and Schröder consider semantic and pragmatic violations, reasoning that language teachers do the same in understanding ill-formed expressions of students (pp.26-27). They have analyzed the ability of their implemented system to eliminate candidate analyses and arrive at a single indicated analysis and diagnosis. This is not the same as evaluating how well the system diagnoses, since it focuses on obtaining unique analyses, but the latter is a significant component of the sort of system needed.

DeSmedt (1995) reports that his grammar tutoring program was able to detect and correct 92.8% of all errors with a false correction rate of only 1.1%. He attributes this to an alternative parsing technique, more strongly driven by expectation, but his description of the parser makes it sound like a version of Earley parsing (DeSmedt 1995, pp158-59), supplemented by techniques which find closest matches. The results are in any case impressive.

36.7 Evaluation

As Holland and Kaplan (1995, pp.373ff) note, most of the work using NLP in CALL has not been evaluated carefully. They attribute this to the pressures of projects which have to show practical use, leaving little time for analysis. Goodfellow (1999) focuses

on "evaluation", and contains valuable remarks, but it studies a single system and shies from objective measures of performance. Some honorable exceptions to the trend of unevaluated work are noted above in the sections on Morphology and Grammar.

Both psycholinguist Brian MacWhinney and second-language learning expert Nina Garrett identify large ranges of questions that need to be asked (MacWhinney 1995, Garrett 1995), even if one would hope that some of the issues MacWhinney raises could be obviated through the accumulated expertise of evaluation in (non-computer-assisted) second-language learning (e.g., issues about which controls and comparisons should be most useful). Certainly projects should plan quantitative evaluation as part of their work.

36.8 Conspectus and Prospectus

Whither research in NLP-inspired CALL?

Given the extensive industrial development of CALL packages, surely Cameron (1999*a*, p.5) is sensible in advocating further cooperation for researchers. And Zock's (1998) plea for interdisciplinary collaboration with linguists and pedagogues bears repetition: many of the issues that arise in constructing and evaluating systems require the expertise of second-language learning experts.

There are many reasons why NLP-based CALL ought to be an attractive area of research for NLP groups. As noted in Sec. 36.2.2 above, there is genuine, substantial demand for language learning services that current supply is unlikely to be able to satisfy (at least not cheaply and generally). As the Sections 36.6.2–36.6.6 demonstrate, there is opportunity to apply a wide range of technologies. And, as noted in Sec. 36.6.4 above, users are intelligent but undemanding—so that technologies which still function less than 100% technically can provide a useful and appreciated service.

The opportunities for NLP collaboration with language learning experts should be optimal in view of their widespread positive reactive to Levy's (1997) book with its plea that the field should emphasize tools rather tutoring systems, and its conclusion that a practical focus would be salubrious for CALL.

All of these factors argue that that a serious investment in this application area is

promising, including the investment in serious evaluation, in incremental improvement, and the investment in collaborations needed to make use of all available expertise.

Further reading and relevant resources

For Computational Linguists interested in CALL, the introduction in Cameron (1999*b*) and the other articles in that collection are probably the best starting place on CALL, and Holland, Kaplan and Sams (1995) the only good collection on opportunities for NLP (and fortunately, it is very good). Cameron (1999*b*) contains excellent articles on several subfields of CALL, in particular on the use of speech technology Aist), parsing (Kempen), grammar checking (Tschichold), corpora (Holmes), and intelligent tutoring systems (Hamburger). Although I was critical of some aspects of Levy (1997) above, it is strong on the history of the field and also the concerns of language educators. Pennington (1996) and Ellis (1994) are older, but contain further useful papers. Holland, Kaplan and Sams (1995) is the only longer work with extensive presentation of the possibilities for NLP in CALL since Last (1989), now dated. The collections cited in Section 36.5.2 comprise a great deal of what's been done recently in NLP applied to CALL.

There are several professional societies devoted to promoting Computer-Assisted Language Learning. **CALICO**, the Computer Assisted Language Instruction Consortium, is a professional organization with an emphasis on modern language teaching and learning. The oldest international group devoted to CALL, it sponsors an annual conference, several special interest groups, and a discussion list. See `calico.org` for a review site for software packages, and an online bibliography. CALICO also hosts **CALICO-L**, an active discussion list at the CALICO web site.

EUROCALL, a large professional organization, has held regular conferences since 1988. It has published the journal *ReCALL Journal* since 1989. See `www.eurocall.org`. **IALLT**, The International Association for Language Learning Technology, publishes the journal *The IALL Journal of Language Learning Technologies*, which was *The Journal of Educational Techniques and Technologies* from 1987-89, and earlier (1969-86) the *NALLD Journal*, i.e., newsletter of the National Association of Language Laboratory Directors. See `iallt.org`.

In addition to the journals published by the professional societies immedaitely above, there are several more journals devoted substantially to CALL and related topics. **ALSIC**, *Apprentissage des langues et systèmes d'information et de communication*, is an journal edited by Thierry Chanier with a French language focus and distinct tracks for research articles and reports on practice. **LLT**, *Language Learning & Technology*, is a refereed electronic journal first published in 1997 and available at `llt.msu.edu`. **System** is a journal on education technology which regularly carries CALL articles.

Syracuse Language Systems (see `www.syrlang.com`) is perhaps the leading company marketing CALL, and has been innovative in including speech and NLP technology. **Agora** (`www.agoralang.com`) maintains a hub of pointers to (mostly commercial) language learning products and services.

Finally, there are two interesting resources which are harder to classify. **NTLL**, "New Technologies and Language Learning", is a **Thematic Network Project** (TNP) in the Area of Languages, coordinated by the *Freie Universität Berlin*. The TNP is closely affiliated with the **European Language Council**, an independent European association for the quantitative and qualitative improvement of knowledge of the languages and cultures of the European Union and beyond. See `www.let.rug.nl/projects/tnp-ntll/`. And **History of CALL** is a web site with lots of pointers concerning the history of CALL, displays with some infelicity in Netscape (in 3/2001), but worthwhile for the content. See `www.history-of-call.org`36. *

References

Aist, G. 1999. "Speech Recognition in Computer-Assisted Language Learning." In *Computer-Assisted Language Learning: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets pp. 165–181.

Allen, J. R. 1996/1997. "Ten Desiderata for Computer-Assisted Language Learning Programms." *Computers and the Humanties* 30(6):441–455.

Antworth, E. L. 1992. "Glossing Text with the PC-KIMMO Morphological Parser." *Computers and the Humanities* 26(5-6):389–98.

Appel, C. and T. Mullen. 2000*a*. "A Common Gateway Interface for Tandem Language Learning." In *Proceedings of the International Congress on Technology in Teaching, jointly organized by IATEFL Computer SIG, TESOL-Spain and Universidad Europea de Madrid, 20-22 Nov. 1998*, ed. B. Strotmann. Universidad Europea: CEES Ediciones.

Appel, C. and T. Mullen. 2000*b*. "Pedagogical Considerations for a Web-Based Tandem Language Learning Environment." *Computers & Education* 34(3–4):291–308.

Appelo, L. and F. M. de Jong, eds. 1994. *Computer-Assisted Language Learning: Proc. of the Seventh Twente Workshop on Language Technology*. Twente: Parlevink.

Atwell, E. 1998. "What can SALT offer the English Teaching Professional?" *English Teaching Professional* 7:46–47.

Atwell, E. 1999. *The Language Machine: The Impact of Speech and Langauge Technologies on English Language Teaching*. London: British Council. avail. online `http://www.comp.leeds.ac.uk/eric`.

Barlow, M. 1996. "Parallel Texts in Language Teaching." In *Proceedings of Teaching and Language Corpora*, ed. S. Botley, J. Glass and T. M. A. Wilson. Number 9 *in* "UCREL Technical Papers" pp. 45–56.

Bauer, D., F. Segond and A. Zaenen. 1995. "LOCOLEX: Translation Rolls off Your Tongue." In *Conference Abstracts: ACH-ALLC '95*. Santa Barbara, USA: . avail. at `http://www.xrce.xerox.com/publis/mltt/mlttart.html`.

Bernstein, J., J. D. Jong, D. Pison and B. Townshend. 2000. "Two Experiments on Automatic Scoring of Spoken Language Proficiency." In *Integrating Speech Technology in the (Language) Learning and Assistive Interface*. Dundee, Scotland: The University of Abertay Dundee pp. 57–61. Prepublication of papers from the Workshop on Speech in CALL at EuroCALL 2000.

Breidt, E. and H. Feldweg. 1997. "Accessing Foreign Languages with COMPASS." *Machine Translation* 12(1-2):153–174. spec. iss. on New Tools for Human Translators.

Cameron, K. 1999*a*. "Introduction." In *Computer-Assisted Language Learning: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets & Zeitlinger.

Cameron, K., ed. 1999*b*. *Computer-Assisted Language Learning: Media, Design and Applications*. Lisse: Swets & Zeitlinger.

Carlson, R., C. Dunger, B. Granstrom and A. Oster, eds. 1998. *STiLL—Speech Technology in Language Learning*. Stockholm: ISCA.

Carter, R. [1]1987, 1998. *Vocabulary: Applied Linguistic Perspectives. 2nd Ed.* London: Routledge.

Connolly, J. H. 1997. "Quantifying Target-Realization Differences." *Clinical Linguistics and Phonetics* 11:267–298.

DeKeyser, R. 1995. "Learning Second Language Grammar Rules: An Experiment with a Miniature Linguistic System." *Studies in Second Language Acquisition* 17(3):379–410.

DeKeyser, R. 2000. "The Robustness of Critical Period Effects in Second Language Acquisition." *Studies in Second Language Acquisition* 22(4):in press.

DeSmedt, W. 1995. "Herr Kommissar: An ICALL Conversation Simulator for Intermediate German." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 153–174.

Dokter, D. and J. Nerbonne. 1998. "A Session with Glosser-R*u*G." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 89–95.

Dokter, D., J. Nerbonne, L. Schurcks-Grozeva and P. Smit. 1998. "Glosser-R*u*G: A User Study." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 169–178.

Dorr, B. 1997. "Large-Scale Acquisition of LCS-Based Lexicons for Foreign Language Tutoring." In *The Fifth Conference on Applied Natural Language*. ACL pp. 139–146.

Dulay, H., M. Burt and S. Krashen. 1982. *Language Two*. Oxford: Oxford University Press.

Ellis, R. 1994. *The Study of Second-Language Acquisition*. Oxford: Oxford University Press.

Felshin, S. 1995. "The Athena Language Learning Project NLP System: A Multilingual System for Conversation-Based Language Learning." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 257–272.

Fox, D. 1996. Final Market Study Report. RECALL (LE 1615) D-28 Exlingua International & Linguistic Engineering of DG XIII (Luxemburg) Belfast: .

FSI. 1973. Expected Levels of Absolute Speaking Proficiency in Languages Taught at the Foreign Service Institute. Technical report Foreign Service Institute, School of Language Studies Monterey, CA: .

Garrett, N. 1995. "ICALL and Second Language Acquisition." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 345–358.

Goethals, M., L. Engels and T. Leenders. 1990. "Automated Analysis of the Vocabulary of English Texts and Generation of Practice Materials: From Main Frame to PC, the Way to the Teacher's Electronic Desk." In *Learning, Keeping, and Using Language: Selected Papers from the Eighth World Congress of Applied Linguistics*, ed. M. Halliday, J. Gibbons and H. Nicholas. Amsterdam and Philadelphia: John Benjamins pp. 231–268.

Goodfellow, R. 1999. "Evaluating Performance, Approach and Outcome." In *Computer-Assisted Language Learning: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets & Zeitlinger pp. 109–140.

Hamburger, H. 1995. "Tutorial Tools for Language Learning by Two-Medium Dialogue." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 183–199.

Hamilton, S. 1998. "RECALL—Some Implications of Learner as User in CALL." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 200–208.

Hoffman, R. 1996. "Computer Networks: Webs of Communication For Language Teaching." In *The Power of CALL*, ed. M. C. Pennington. Houston: Athelstan.

Holland, V., ed. 1999. *Tutors that Listen: Speech Recognition for Language Learning.* Special Issue of *CALICO* 16(3).

Holland, V. M. 1995. "The Case for Intelligent CALL." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. vii–xvi.

Holland, V. M. and J. D. Kaplan. 1995. "Natural Language Processing Techniques in Computer Assisted Language Learning: Status and Instructional Issues." *Instructional Science* 23:351–380.

Holland, V. M., J. D. Kaplan and M. R. Sams, eds. 1995. *Intelligent Language Tutors: Theory Shaping Technology.* Mahwah, NJ: Lawrence Erlbaum.

Holliday, L. 1998. "The Grammar of Second Language Learners of English EMAIL Messages." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 136–145.

Holmes, G. 1999. "Corpus CALL: Corpora in Language and Literature." In *Computer-Assisted Language Learning: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets pp. 239–269.

Hulstijn, J. 1992. "Retention of Inferred and Given Word Meanings: Experiments in Incidental Vocabulary Learning." In *Vocabulary and Applied Linguistics*, ed. P. J.L.Arnaud and H. Béjoint. London: Macmillan pp. 113–125.

Jager, S. 1998. "Hologram—Computer-Assisted Academic Grammar Learning." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 82–88.

Jager, S., J. Nerbonne and A. van Essen, eds. 1998. *Language Teaching and Language Technology*. Lisse: Swets and Zeitlinger.

Jensen, K., G. Heidorn, L. Miller and Y. Ravin. 1993. "Prose Fitting and Prose Fixing." In *Natural Language Processing: The PNLP Approach*, ed. K. Jensen, G. Heidorn and S.D.Richardson. Boston: Kluwer pp. 53–64.

Johns, T. and P. King. 1991. "Should you be Persuaded: Two Samples of Data-Driven Learning Materials." In *Classroom Concordancing*, ed. T. Johns and P. King. Birmingham: Birmingham University pp. 1–16. English Language Research Journal 4.

Kempen, G. 1992. "Language Technology and Language Instruction: Computational Diagnosis of Word-Level Errors." In *Intelligent Tutoring Systems for Foreign Language Learning*, ed. M.L.Swartz and M.Yazdani. Berlin: Springer.

Kempen, G. 1999. "Visual Grammar: Multimedia for Grammar and Spelling Instruction in Primary Education." In *CALL: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets and Zeitlinger pp. 223–238.

Kempen, G. and A. Dijkstra. 1994. "Toward an Integrated System for Grammar, Writing, and Spelling Instruction." In *Computer-Assisted Language Learning: Proc. of the Seventh Twente Workshop on Language Technology*, ed. L. Appelo and F. M. de Jong. Twente: Parlevink.

Krantz, G. 1991. "Learning Vocabulary in a Foreign Language" PhD thesis Göteborg.

Krashen, S. D. 1982. "The Fundamental Pedagogical Principle in Second-Language Teaching." *Studia Linguistica* 35(1–2):50–71.

Krashen, S. D. 1985. *The Input Hypothesis*. London: Longman.

Krashen, S. D. 1989. "We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis." *Modern Language Journal* 73(4):440–464.

Lantolf, J. P. 1996. "SLA Theory Building: "Letting All the Flowers Bloom!"." *Language Learning* 46(4):713–49.

Larsen-Freeman, D. and M. Long. 1991. *An Introduction to Second Language Acquisition Research.* London: Longman.

Last, R. 1989. *Artificial Intelligence Techniques in Language Learning.* Chichester: Horwood.

Levin, L. S. 1993. "Translators' Workstations and Language Learners' Workstations." In *Progress in Machine Translation*, ed. S. Nirenburg. Amsterdam: IOS Press pp. 287–290.

Levin, L. S. and D. A. Evans. 1995. "ALICE-chan: A Case Study in ICALL Theory and Practice." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 77–98.

Levy, M. 1997. *Computer-Assisted Language Learning: Context and Conceptualization.* Oxford: Oxford University Press.

Loritz, D. 1995. "GPARS: A Suite of Grammar Assessment Systems." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 121–134.

MacWhinney, B. 1995. "Evaluating Foreign Language Tutoring Systems." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 317–326.

McCoy, K. F., C. A. Pennington and L. Z. Suri. 1966. "English Error Correction: A Syntactic User Model Based on Principled 'Mal-Rule' Scoring." In *Proceedings of User-Modeling 1996, The Fifth International Conference on User Modeling.* Kailua-Kona, Hawaii: pp. 59–66.

Menzel, W. and I. Schröder. 1999. "Error Diagnosis for Language Learning Systems." *ReCALL* Spec.Publication:20–30. Language Processing in CALL.

Michaud, L. and K. McCoy. 2000. "Supporting Intelligent Tutoring in CALL by Modeling the User's Grammar." In *Proc. of the* 13^{th} *Annual International Florida Artificial Intelligence Research Symposium, Special Track on AI in Instructional Soft-*

ware. Orlando, Fla: American Association for Artificial Intelligence. avail. at `http://www.eecis.udel.edu/research/icicle/pubs/MichMcCo00.ps`.

Miller, J. F. and T. Klee. 1995. "Computational Approaches to the Analysis of Language Impairment." In *The Handbook of Child Language*, ed. P. Fletcher and B. MacWhinney. Oxford: Blackwell pp. 545–72.

Mondria, J.-A. 1996. "Vocabulaireverwerving in het vreemde-talenonderwijs: De effecten van context en raden op de retentie" PhD thesis University of Groningen Groningen, The Netherlands: .

Murphy, M., A. Krüger and A. Grieszl. 1998. "RECALL—Providing an Individualized CALL Environment." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 62–73.

Murray, J. H. 1995. "Lessons Learned from the Athena Language Learning Project: Using Natural Language Processing, Graphics, Speech, and Interactive Video for Communication-Based Language Learning." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 243–256.

Nerbonne, J. 2000. "Parallel Texts in Computer-Assisted Language Learning." In *Parallel Text Processing*, ed. J. Veronis. Dordrecht: Kluwer pp. 354–369.

Nerbonne, J. and D. Dokter. 1999. "An Intelligent Word-Based Language Learning Assistant." *Traitment Automatique des Languages* 40(1):125–142. Special issue on Multilingual Processing edited by Remi Zajac.

Nerbonne, J., D. Dokter and P. Smit. 1998. "Morphological Processing and Computer-Assisted Language Learning." *Computer-Assisted Language Learning* 11(5):421–37. Special issue on French edited by Michael Zock.

Nerbonne, J., L. Karttunen, E. Paskaleva, G. Prószéky and T. Roosmaa. 1997. "Reading More into Foreign Languages." In *The Fifth Conference on Applied Natural Language*. ACL pp. 135–138.

Nerbonne, J., P. Kleiweg, P. Smit, E. Kuipers and D. Dokter. 2001. "A Web-Based Foreign Language Assistant." In *Essays in Language, Language Education, and the History of Linguistics in Honour of Arthur van Essen.* Amsterdam and Philadelphia: John Benjamins. avail. at `www.let.rug.nl/ nerbonne/papers/gloss-web/`.

Nerbonne, J., S. Jager and A. van Essen. 1998. "Language Teaching and Language Technology: An Introduction." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 1–10.

Norris, J. and L. Ortega. 2000. "Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-Analysis." *Language Learning* 50(3):in press.

Olsen, M. B., ed. 1999. *Computer-Mediated Language Assessment and Evaluation in Natural Language Processing.* College Park, Maryland: ACL.

Pasero, R. and P. Sabatier. 1998. "Linguistic Games for Language Learning." *Computer-Assisted Language Learning* 11(5):561–585. Special issue on French edited by Michael Zock.

Paskaleva, E. and S. Mihov. 1998. "Second Language Acquisition from Aligned Corpora." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 43–52.

Pennington, M. C. 1996. *The Power of Computer-Assisted Language Learning.* Houston: Athelstan.

Pennington, M. C. 1999. "The Missing Link in Computer-Assisted Writing." In *Computer-Assisted Language Learning: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets & Zeitlinger pp. 271–293.

Perry, C. K. 1995. "Review of Phonological Deviation Analysis by Computer (PDAC)." *Child Language Teaching and Therapy* 11:331–340.

Peters, C., E. Picchi and L. Biagini. 1996. "Parallel and Comparable Bilingual Corpora in Language Teaching and Learning." In *Proceedings of Teaching and Language*

Corpora, ed. S. Botley, J. Glass and T. M. A. Wilson. Number 9 *in* "UCREL Technical Papers" pp. 68–82.

Proceedings of the Fifth Conference on Applied Natural Language Processing. 1997. Washington: Association for Compuational Linguistics. with spec. section on CALL, pp.127–146.

Roosmaa, T. and G. Proszeky. 1998. "GLOSSER—Using Language Technology Tools for Reading Texts in a Foreign Language." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 101–107.

Rothenberg, M. 1998. "The New Face of Distance Learning in Language Instruction." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 146–48.

Rypa, M. and K. Feuerman. 1995. "CALLE: An Exploratory Environment for Foreign Language Learning." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 55–76.

Salaberry, M. R. 1996. "A Theoretical Foundation for the Development of Pedagogical Tasks in Computer-Mediated Communication." *CALICO Journal* 14(1):5–34.

Salaberry, M. R. 2000. "Review of S.Jager, J.Nerbonne & A. van Essen *Language Teaching and Language Technology*." *Language Learning & Technology* 4(1):22–25. avail. online at `http://llt.msu.edu/`.

Sanders, R. H. and A. F. Sanders. 1995. "History of an AI Spy Game: Spion." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 141–152.

Schneider, D. and K. McCoy. 1998. "Recognizing Syntactic Errors in the Writing of Second Language Learners." In *Proceedings of the* 36^{th} *Annual Meeting of the ACL and the* 17^{th} *COLING*. Montreal: ACL pp. 1198–1204.

Schoelles, M. and H. Hamburger. 1997. "The NLP Role in Animated Conversation for CALL." In *The Fifth Conference on Applied Natural Language*. ACL pp. 127–134.

Schultz, J.-M. 2000. "Computers and Collaborative Writing in the Foreign Language Classroom." In *Network-based Language Teaching: Concepts and Practice:Concepts and Practice*, ed. M. Warschauer and R. Kern. Cambridge: Cambridge University Press.

Schulze, M. and M.-J. Hamel. 1998. "Conference Report on *Natural Language Processing in Computer-Assistend Language Learning*, Centre for Computational Linguistics, UMIST, 9 May 1998." *ReCALL Journal* 10(2):55–56. avail. at `http://www.eurocall.org/recall/`.

Schulze, M., M.-J. Hamel and J. Thompson, eds. 1999. *Language Processing in CALL*. Hull: CTICML. *ReCALL* Special Publication (Proceedings of a one- day conference "Natural Language Processing in Computer-Assisted Language Learning" held at UMIST, 9 May 1998, organised by the Centre of Computational Linguistics, UMIST, in association with EUROCALL).

Somers, H. 1999. "Aligning phonetic Segments for Children's Articulation Assessment." *Computational Linguistics* 25(2):267–275.

Sparck Jones, K. 1991. "Tailoring Output to the User: What does User Modeling in Generation Mean?" In *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, ed. C. Paris, W. Swartout and W. Mann. Boston: Kluwer Academic pp. 201–225.

Tschichold, C. 1999. "Grammar Checking for CALL: Strategies for Improving Foreign Language Grammar Checkers." In *CALL: Media, Design and Applications*, ed. K. Cameron. Lisse: Swets and Zeitlinger pp. 203–221.

van Els, T., G. Extra, C. van Os and T. Bongaerts. 1977. *Handboek voor de Toegepaste Taalkunde*. Groningen: Wolters-Noordhof. English translation: *Applied Linguistics and the Learning and Teaching of Foreign Languages*, (1984), Edward Arnold, London.

van Heuven, V. 1998. "COOL/CALP: Computer-Assisted Learning to Parse in Dutch." In *Language Teaching and Language Technology*, ed. S. Jager, J. Nerbonne and A. van Essen. Lisse: Swets and Zeitlinger pp. 74–81.

Warschauer, M. 1996. "Computer-Assisted Language Learning: An Introduction." In *Multimedia language teaching*, ed. S. Fotos. Tokyo: Logos International pp. 3–20.

Weinberg, A., J. Garman, J. Martin and P. Merlo. 1995. "A Principle-Based Parser for Foreign Language Tutoring in German and Arabic." In *Intelligent Language Tutors: Theory Shaping Technology*, ed. V. M. Holland, J. D. Kaplan and M. R. Sams. Mahwah, NJ: Lawrence Erlbaum pp. 23–44.

Weischedel, R., W. Voge and M. James. 1978. "An Artificial Intelligence Approach to Language Instruction." *Artificial Intelligence* 10:225–240.

Wichmann, A., S. Fligelstone, T. McEnery and G. Knowles. 1997. *Teaching and Language Corpora*. New York: Addison-Wesley Longman.

Widdowson, H. 1990. *Aspects of Language Teaching*. Oxford: Oxford University Press.

Wolff, D. 1999. "Review of M.Levy *Computer-Assisted Language Learning: Context and Conceptualization*." *System* 27(1):125–128.

Zock, M. 1992. "SWIM or SINK: the Problem of Communicating Thought." In *Bridge to International Communication: Intelligent Tutoring Systems for Foreign Language Learning*, ed. M.L.Swartz and M.Yazdani. Berlin: Springer pp. 235–247.

Zock, M. 1994. "Language in Action, or, Learning a Language by Seeing it Work." In *Computer-Assisted Language Learning: Proc. of the Seventh Twente Workshop on Language Technology*, ed. L. Appelo and F. M. de Jong. Twente: Parlevink.

Zock, M. 1996. "Computational Linguistics and its Use in the Real World: The Case of Computer-Assisted Language Learning." In *Proc. of the 16th COLING*. Copenhagen: ICCL pp. 1002–1004.

Zock, M. 1998. "Guest Editorial Introducing Special Issue on French." *Computer-Assisted Language Learning* 11(5):467–474.

Zock, M., A. Laroui and G.Francopoulo. 1989. "SWIM: A 'Natural' Interface for the Scientifically Minded Language Learner." *Computers and the Humanities* 23:411–422.