

The Computer and Natural Language (Ling 445/515)

Topic 7: Computer-Aided Language Learning

Markus Dickinson
Dept. of Linguistics, Indiana
Autumn 2008

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

1 / 52

Some common computer uses

- ▶ Computers are widely used in support of foreign language teaching (FLT). For example, they
 - ▶ provide access to foreign language newspapers, radio, and TV programs through the world-wide web
 - ▶ connect language learners with native speakers through email/chat
 - ▶ support multimedia presentations providing an audio-visual foreign language context
 - ▶ enable the learner to search for real-life examples in electronic corpora
- ▶ Essentially, such computer usage helps language learners experience a foreign language and culture in a more direct, real-life fashion.

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

2 / 52

Awareness of language forms is important

- ▶ Which other aspects of language learning are relevant in this context, apart from the undisputed importance of contextualized, communicative language usage?
- ▶ Research since the 90s has shown that **awareness** of language forms and rules is important for an adult learner to successfully acquire a foreign language.
 - ▶ CALL can provide an opportunity to enhance awareness of a language's rules

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

3 / 52

Real-life constraints

- ▶ The time a student can spend with an instructor/tutor typically is very limited.
- ▶ In consequence, work on form and grammar is often deemphasized and confined to homework so that the time with the instructor can be used for purely communicative activities.
- ▶ The downside is that the learner has relatively few opportunities to gain awareness of forms and rules and receive individual feedback on errors.

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

4 / 52

An opportunity for CALL

- ▶ The situation seems like an excellent opportunity for developing Computer-Aided Language Learning (CALL) tools to
 - ▶ provide individual feedback on learner errors and
 - ▶ foster learner awareness of relevant language forms and categories.
- ▶ But existing CALL systems which offer exercises
 - ▶ typically are limited to uncontextualized multiple choice, point-and-click, or simple form filling, and
 - ▶ feedback usually is limited to yes/no or letter-by-letter matching of the string with a pre-stored answer.
 - ▶ An example for letter-by-letter feedback on the "Spanish Grammar Exercises" site (B. K. Nelson).

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

5 / 52

From CALL to ICALL

- ▶ Linguistic modeling is needed to improve on this situation, e.g.:
 - ▶ tokenization: identify words
 - ▶ morphological analysis: identify/interpret morphemes
 - ▶ syntactic analysis: identify selection, government and agreement relations and word order requirements
 - ▶ formal pragmatic analysis: identify coreference relations, information structure partitioning, . . .
- ▶ Computational tools identifying such linguistic properties need to be integrated into CALL systems to obtain language-aware "Intelligent" CALL (ICALL).
 - ▶ Tools must be extended/written to permit and diagnose errors made by language learners.

CALL is a big business: 106 million Euro (about \$120 million) spent on CALL products in Europe in 1994, and US market is twice as big. (Nerbonne 2003)

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

6 / 52

Expectations

How successful we think CALL systems are depends upon our expectations and upon the expectations of the students

- ▶ Unlike classroom situations, where the instructor can be assumed to be right, ICALL systems are prone to errors
 - ▶ Parsers return wrong analyses: students could be told they are right when they are wrong, and vice versa
 - ▶ Parsers are incomplete: no analysis may be returned
- ▶ Students have to understand up front that this is not a human
 - ▶ Perhaps obvious, but affects, for example, the image of a tutor you might have on screen

Computers and Language

Topic 7: CALL

What is ICALL?

Awareness of language forms

An opportunity for CALL

SLA

Language learning

Error analysis

Feedback

Student model

CALL systems

Basic uses of computers

Early CALL systems

ICALL

NLP uses for CALL

Authentic Text ICALL

Parser-Based ICALL

Parser Issues

7 / 52

Second Language Acquisition (SLA)

If we want to design a system to teach, e.g., English progressives to Japanese speakers, we have to ask:

- ▶ What kinds of exercises are best for learning this construction, or for any sort of grammatical principle?
- ▶ How advanced are these learners? Should we even bother teaching this to novice learners?
- ▶ Do Japanese speakers have trouble with this construction, or is it pretty easy to grasp?
- ▶ Is a given individual generally a good student? Do they have problems with certain constructions?

Insights from research on **second language acquisition (SLA)** can provide a starting point to answer these questions

Computers and Language

Topic 7: CALL

What is ICALL?

Awareness of language forms

An opportunity for CALL

SLA

Language learning

Error analysis

Feedback

Student model

CALL systems

Basic uses of computers

Early CALL systems

ICALL

NLP uses for CALL

Authentic Text ICALL

Parser-Based ICALL

Parser Issues

8 / 52

Explicit instruction

There are crucial differences between:

- ▶ First language acquisition (L1), and
- ▶ **Second language acquisition (L2)**

For first language acquisition,

- ▶ There is less explicit instruction:
 - ▶ No one tells babies what the rules of language are, but ...
 - ▶ Motherese ("baby talk"): different lexicon, intonation, topics, turn-taking
- ▶ Parents generally correct content, not grammar.
 - ▶ Positive evidence of language: children learn language based on all the possible sentence around them.
- ▶ Babies tune in on relevant distinctions: synapses still taking shape in the brain

Computers and Language

Topic 7: CALL

What is ICALL?

Awareness of language forms

An opportunity for CALL

SLA

Language learning

Error analysis

Feedback

Student model

CALL systems

Basic uses of computers

Early CALL systems

ICALL

NLP uses for CALL

Authentic Text ICALL

Parser-Based ICALL

Parser Issues

9 / 52

Stages of Adult Language Learning

1. Silent/Preproduction Stage:
 - ▶ about 500 words
2. Early Production Stage:
 - ▶ about 1000 active words
 - ▶ short utterances
3. Speech Emergence Stage:
 - ▶ about 3000 words
 - ▶ short phrases, simple sentences, first dialogues
4. Intermediate Language Proficiency Stage:
 - ▶ about 6000 words
 - ▶ complex statements, state opinions/thoughts
5. Advanced Language Proficiency Stage
 - ▶ extensive, specialized vocabulary
 - ▶ full dialogues
 - ▶ can take 5-7 years to get here.

Knowing stage is important for developing CALL systems

Computers and Language

Topic 7: CALL

What is ICALL?

Awareness of language forms

An opportunity for CALL

SLA

Language learning

Error analysis

Feedback

Student model

CALL systems

Basic uses of computers

Early CALL systems

ICALL

NLP uses for CALL

Authentic Text ICALL

Parser-Based ICALL

Parser Issues

10 / 52

Language-specific stages of learning

Those are general learning stages

- ▶ It can also help to identify when learners typically pick up specific constructions.
- ▶ English learners, for example, typically follow this pattern:
 1. Progressive *-ing*, plurals, the verb *to be*
 2. Auxiliary verbs, articles (*a, an, the*)
 3. Irregular past tense
 4. Regular past tense, third person singular *-s*, possessive *-s*

Computers and Language

Topic 7: CALL

What is ICALL?

Awareness of language forms

An opportunity for CALL

SLA

Language learning

Error analysis

Feedback

Student model

CALL systems

Basic uses of computers

Early CALL systems

ICALL

NLP uses for CALL

Authentic Text ICALL

Parser-Based ICALL

Parser Issues

11 / 52

Error analysis

- ▶ Given multiple possible analyses for a sentence, which one is most likely, based on:
 - ▶ the stage of acquisition of the learner
 - ▶ the first language of the learner
 - ▶ the focus of the exercise
- ▶ What kinds of errors are we interested in and do we expect?
 - ▶ We would like an **error typology** = a classification of errors into different groups.
 - ▶ People can look at error-tagged corpora to find the most common errors

Computers and Language

Topic 7: CALL

What is ICALL?

Awareness of language forms

An opportunity for CALL

SLA

Language learning

Error analysis

Feedback

Student model

CALL systems

Basic uses of computers

Early CALL systems

ICALL

NLP uses for CALL

Authentic Text ICALL

Parser-Based ICALL

Parser Issues

12 / 52

Feedback

Feedback = response to the learner based on their input.

Purpose of feedback:

- ▶ Reinforcement: feedback can act as a reinforcer to learn a particular concept (behaviorism)
- ▶ Learning processes need feedback to know right from wrong (cognitivism)

Things to keep in mind when designing a system (somewhat obvious):

- ▶ Feedback needs to be accurate.
- ▶ Displaying more than one error message at a time is not helpful (Heift 2001).
- ▶ Explanations should be short.

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis

Feedback

Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

13 / 52

Kinds of feedback (Schulze 2003)

Assume a student incorrectly says *Yo habla español*

- ▶ **Explicit correction** = explicitly giving the correct form, indicating that this is a correction.

TEACHER: No, not *Yo habla*. You want to say *Yo hablo*.

- ▶ **Recast** = reformulating all or part of the student's utterance, without the error, and not indicating that this is a correction.

STUDENT: *Yo habla español*.

TEACHER: *Yo hablo español* tambien. (I speak Spanish also.)

- ▶ **Clarification request** = asking for a clarification.

TEACHER: What? What did you mean?

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis

Feedback

Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

14 / 52

Kinds of feedback (cont.)

- ▶ **Metalinguistic feedback** = comments, information, or questions about the correctness of the student's utterance, but not giving the correct form.

TEACHER: Now, if you're speaking in the first person, is that the verb form you want?

- ▶ **Elicitation** = eliciting a response from a student by pausing or by asking the student to reformulate the response.

STUDENT: *Yo habla*.

TEACHER: Mm-hmm. ...

STUDENT: *Yo hablo*.

- ▶ **Repetition** = repeating the student's utterance, usually marking the error intentionally.

TEACHER: *Yo habla?*

SLA has a lot to say about what feedback is best (but not necessarily in the context of ICALL)

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis

Feedback

Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

15 / 52

Student model

Often, when focusing on NLP, we keep a student model

- ▶ Allows us to say which grammatical points need more work for a given individual.
- ▶ Allows us to give different feedback and/or exercises based on the learner's abilities.
 - ▶ Language abilities
 - ▶ Cognitive reasoning abilities
 - ▶ General mental abilities, e.g., short-term memory
- ▶ Ensures that the learner knows the terminology presented in the feedback

Student modeling faces questions like:

- ▶ Should the student have access to their model and be able to self-update it?
- ▶ How should the student information be integrated into the rest of the system?

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis

Feedback

Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

16 / 52

CALL Systems

Having discussed what we need to know when learning a language, we can now focus on:

- ▶ Basic uses of computers for CALL
- ▶ Early CALL Systems
- ▶ Intelligent CALL (ICALL)
 - ▶ Straightforward uses of NLP for CALL
 - ▶ Authentic Text ICALL Systems
 - ▶ Parser-Based ICALL Systems

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis

Feedback

Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

17 / 52

Basic uses of computers for CALL

Lots of general possibilities for using a computer to learn:

- ▶ multimedia presentations
- ▶ online dictionaries with fast access
- ▶ extensive databases of information
- ▶ digital audio files
- ▶ digital videos of people speaking in L2
 - ▶ Digital advantages: easy playback, easy isolation of problematic spots, etc.
- ▶ interactive games & puzzles
- ▶ exercises for students to complete

The last two examples potentially require sophisticated natural language processing

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis

Feedback

Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

18 / 52

Early CALL systems

Some older types of systems are along the following lines

- ▶ **Frame-based systems** “match student answers with a set of correct and incorrect answers stored in a frame”

These systems differ in their strategies for selecting questions, but they rely on preset questions and answers

- ▶ In principle, could be used with NLP techniques

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

19 / 52

Linear systems

A **linear system** does the following:

1. poses a question
2. accepts an answer
3. informs the student as to whether or not the answer was correct

⇒ Regardless of the correctness of the answer, linear systems proceed to the next question.

- ▶ **Branching systems** select a question of appropriate difficulty based on performance thus far
- ▶ **Generative systems** select questions so as to have a different session each time

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

20 / 52

Problems with frame-based systems

Frame-based systems are fairly simple and generally do not involve much linguistic knowledge

- ▶ There is no deep understanding of question domain
- ▶ They merely match answers with questions, but language use is more varied
 - ▶ Could be more than one correct answer, as with a translation task
 - ▶ Fill-in-the-blank exercises (pointed out by Trude Heift):

(1) Today is November 10. What date is tomorrow? Tomorrow is _____.

- a. The eleventh. [Judged correct]
- b. November 11. [Judged incorrect]

- ▶ There is not much tailoring to particular student needs

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

21 / 52

Intelligent CALL (ICALL)

Intelligent CALL (ICALL) focuses on using linguistics and natural language processing to make CALL better.

- ▶ Some basic NLP uses for CALL (Nerbonne 2003):
 - ▶ Concordancers
 - ▶ Text alignment
 - ▶ Speech recognition and synthesis
 - ▶ Morphological processing (for vocabulary help)
- ▶ ICALL can also involve integrating authentic text into exercises, usually for more advanced learners
- ▶ ICALL involves using a parser (syntactic processor) to handle real learner input and deals with issues like:
 - ▶ Diagnosing and accounting for user errors
 - ▶ Modeling the system on particular (kinds of) users
 - ▶ Presenting useful feedback to the user

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

22 / 52

Concordancers

- ▶ Take a text and create a **concordance** = display of words in context.
- ▶ Concordancers help learners understand how a given word is used.
 - ▶ For example, is the word *data* in English singular or plural?

contract to supply voice and giving control over how much humanists to fit their special 27 mm . But these **data** communications within the Tunnel in **data** is sent over the network **data** to the software , rather **data** are for fourth-year crabs .

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

23 / 52

Text alignment

Show learners texts which are aligned between two languages.

- ▶ Advanced learners can benefit by seeing how word usage in native language corresponds to usage in L2
- ▶ Beginning learners would probably be overwhelmed

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

24 / 52

Speech recognition and synthesis

Speech technology can help with pronunciation aspects of learning a language

- ▶ ASR: check and improve pronunciation.
- ▶ TTS: generate pronunciations of isolated words.

If you're using a paper dictionary, you have to base your pronunciations on a phonetic transcription of some sort.

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

25 / 52

Morphological analysis

- ▶ **Lemmatization** = extract the lemma, or stem, of a word. (e.g. lemma of *running* is *run*; lemma of *corpora* is *corpus*.)
- ▶ **Morphological generation** = generate different forms of a word based on its lemma and part of speech, or word class.

These processes are used to:

- ▶ help provide drill material for learners
- ▶ facilitate dictionary lookup (which can be very difficult otherwise for “highly inflected” languages – e.g. the lemma of Russian *berut* is *brat*)

GLOSSER, for example, is a system that uses morphological processing to speed up dictionary look-up (100 times faster) (Nerbonne 2003)

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

26 / 52

Authentic Text ICALL

Authentic Text ICALL attempts to connect learners to appropriate naturally-occurring texts

- ▶ Allows students to find examples in target language related to their interests
- ▶ Allows for more exploration and something akin to “immersion”

There are challenges in:

- ▶ Extracting appropriate exercises from the real texts
- ▶ Finding level-appropriate texts

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

27 / 52

The WERTi System

Working with English Real-Texts: An Intelligent Workbook for English

WERTi is an “intelligent automatic workbook, providing an unlimited number of activities designed to foster awareness of English grammatical forms and functions”

- ▶ Learners select a topic which fits their interests
- ▶ Webpages are returned, which learners interact to learn about, e.g., prepositions
 - ▶ Learners can choose to see prepositions in color; click on them; or fill in blanks

Crucially, the exercises are **generated** on the fly

- ▶ Pre-existing NLP technology (e.g., a POS tagger) is used to spot the relevant categories

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

28 / 52

The REAP Project

Reader-Specific Lexical Practice for Improved Reading Comprehension

In the REAP system:

- ▶ Teachers have target vocabulary items
- ▶ REAP finds appropriate texts for learners, based on their individual profile
 - ▶ Learners get individualized vocabulary practice from authentic web texts

There are several challenges in extracting text for reading

- ▶ Each extracted text is analyzed for its “syntactic features, readability, length, and the occurrence of target vocabulary”
- ▶ Information retrieval and statistical NLP techniques are used to find appropriate texts

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

29 / 52

Parser-Based (I)CALL

Parser-Based ICALL systems generally fall along the following lines:

- ▶ System presents the learner with an exercise
- ▶ Learner inputs an answer, possibly with errors, i.e., a potentially **ill-formed** sentence
- ▶ The parser processes this sentence
 - ▶ Identifying where, if at all, it was incorrect
 - ▶ The nature of the error
- ▶ Feedback is then presented to the student

We'll look at three different example systems:

- ▶ e-Tutor (German Tutor): Heift & Nicholson
- ▶ BANZAI/Robo-Sensei: Nagata
- ▶ TAGARELA: Amaral & Meurers

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

30 / 52

e-Tutor (German Tutor)

e-Tutor (Heift & Nicholson 2001) is used at Simon Fraser University to teach German to students; it is:

- ▶ general, i.e., allows for any native language (L1)
- ▶ able to capture different kinds of errors . . . because the exercises are very constrained

Student input is put through the following modules and stops with feedback when the first error is encountered

1. String match: if the input matches a pre-defined correct answer, we know it's good.
 - ▶ Prevents time-consuming analysis for perfect answers
2. Punctuation check: is any punctuation missing?

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

31 / 52

More on system architecture

3. Spell check: run an off-the-shelf spell checker on the input and get the **lemmas**

- ▶ Idea: eliminate the really basic errors.
- ▶ Problem: sometimes a "misspelled" word is a sign of lack of grammatical competence, e.g. *runned* is "misspelled", but it might show a lack of knowledge about the English past tense.

4. Example check: are the right words being used?

5. Missing word check: are any words missing?

6. Extra word check: are any words added?

- ▶ These 3 steps (example, missing word, and extra word checks) all are based on the notion that the exercise has *pre-defined* all the words which are acceptable for this answer.

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

32 / 52

More on system architecture (cont.)

7. Word order check: match the user word order with the correct word order (a big issue in German)

8. Grammar check

- ▶ This is the most complicated part of the process, the one which requires linguistic knowledge (syntax)
- ▶ About 60% of errors make it to this stage.

9. Catch-all: just in case everything else fails

Note:

- ▶ Heift's system works so well because the exercises themselves are constrained, as we will see
- ▶ The approach is very **modular** = each check is an independent program

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

33 / 52

Kinds of exercises

Here are some example exercises from Heift's system, outlined in

<http://zif.spz.tu-darmstadt.de/jg-06-2/beitrag/heift2.htm>

- ▶ Dictation
- ▶ Build a Phrase
- ▶ Which Word is Different
- ▶ Word Order Practice
- ▶ Fill-in-the-Blank
- ▶ Build a Sentence

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

34 / 52

Dictation

Student hears a sentence in German and types it in. They are told if they are correct, and if not, why.

Guten Tag, Trude!

Hören Sie das Diktat. Hören Sie dann einen Satz und schreiben Sie.

Übung 2 von 6 (Satz 2 von 2)

Guten Tag! Mein Name ist Fumiko Kanno.

Achtung! Rechtschreibung bei dem folgenden Wort:

Japan : Japan

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

35 / 52

Dictation (cont.)

Pros:

- ▶ Input is very constrained.
- ▶ Very useful to be able to practice listening by oneself.
- ▶ Won't take up class time.

Cons:

- ▶ Requires multimedia resources.
- ▶ Takes a long time to prepare.
- ▶ Not very contextualized.

Such good and bad points could be given for all of the following exercises

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

36 / 52

Build a Phrase

Build up a complete phrase (e.g., a noun phrase) based on a given picture—in other words, provide your own vocab.

Guten Tag, Trude! Umlaute + ß

Schreiben Sie das Substantiv mit Artikel.

Übung 2 von 10



die Croissants

Prima!

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

37 / 52

Which Word is Different

Given 4 words, pick the one which differs from the others.

Guten Tag, Trude! Umlaute + ß

Was passt nicht? Klicken Sie das Wort und dann "PRÜFEN".

Übung 5 von 10

Sommer

Uhr

Frühling

Winter

Gut gemacht!

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

38 / 52

Word Order Practice

Take all the given words and arrange them into a sentence.

Guten Tag, Trude!

Ziehen Sie die folgenden Wörter in das Textfeld.

Übung 5 von 15

einem oder 'wohnung? einer Haus

Wohnt du in

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

39 / 52

Fill-in-the-Blank

Give the learner a lemma or choice of lemmas, and they have to fill in the blank.

Guten Tag, Trude! Umlaute + ß

Schreiben Sie die fehlenden Wörter.

Übung 3 von 10

Fumiko wohnt in München.

To!!!

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

40 / 52

Build a Sentence

Use all the given words (lemmas) and create a grammatical German sentence.

Guten Tag, Trude! Umlaute + ß

Bilden Sie einen Satz mit den folgenden Wörtern.

Übung 4 von 10

(def. Artikel) / Zeit / laufen.

Der Zeit läuft.

Da ist ein Genusfehler bei dem Subjekt.

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

41 / 52

BANZAI/Robo-Sensei

Robo-Sensei is proprietary software developed for teaching introductory Japanese courses

- ▶ Exercises cover a range of Japanese constructions
 - ▶ Extensive metalinguistic feedback is given to the learner
 - ▶ Noriko Nagata has shown, to some extent, that more difficult constructions are better served with this kind of feedback
- ▶ Hand-written rules cover Japanese morphology and syntax patterns
 - ▶ Comparison is done between the analysis of a learner's input and a target analysis, in order to spot the error(s) and provide effective feedback

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

42 / 52

Advanced learner output here: "There is an error in gender with the subject."

TAGARELA

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

TAGARELA is a system for individualized instruction of Portuguese at Ohio State

- ▶ It features standard exercises, as found in foreign language workbooks
- ▶ NLP processing is used to detect spelling, morphological, syntactic, and semantic errors
- ▶ A student model is kept to track performance and to choose appropriate feedback
 - ▶ An instruction model allows the instructor to state what is important

TAGARELA system overview

Computers and Language
Topic 7: CALL

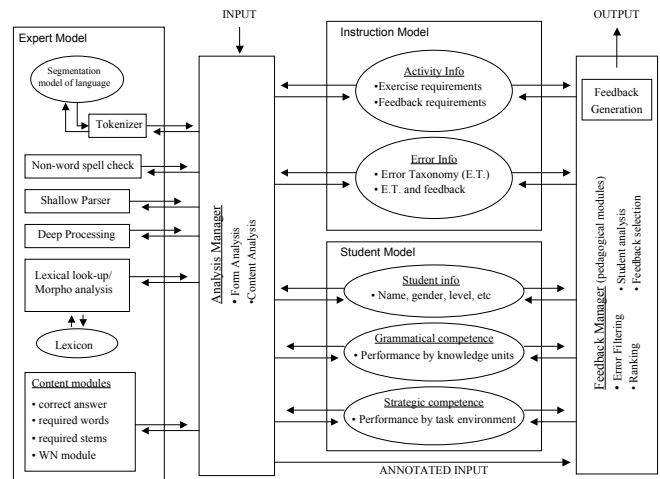
What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues



Demand-driven architecture

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

Different from the e-Tutor, TAGARELA works in a **demand-driven** fashion; the analysis manager:

- ▶ receives input from the student
- ▶ gather the necessary information from:
 - ▶ instruction model
 - ▶ student model
- ▶ decides on the best processing strategy
 - ▶ which NLP modules to call
 - ▶ in which order (as opposed to linearly)
- ▶ calls NLP modules to process input, producing an input annotated with linguistic properties
- ▶ hands the annotated input to the feedback manager

Parser issues

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

The parser-based systems we've just seen have to deal with errors in learner input:

- ▶ How do we adapt our technologies to find errors?
- ▶ Do we tailor the system to a particular kind of learner?
- ▶ Do we tailor the system for an individual learner?
- ▶ What is the exact error?
- ▶ How many errors are there?

We will examine each of these in turn

How do we adapt technology to find errors?

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

Our parsers, morphological analyzers, and so on, are made to handle well-formed input.

- ▶ Use so-called **mal-rules** = rules which are added to your grammar that handle error cases.
 - ▶ e.g., A singular noun and a plural verb are allowed to combine, but it is marked as an error.
 - ▶ $S_{error} \rightarrow NP_{sg} VP_{pl}$
- ▶ Modify your technology: a parser can be reworked to handle ill-formed input.
 - ▶ Parsers normally just "die" when handling bad input.
 - ▶ e.g., I'll parse *John are big*, but I'll tell you that I didn't like it and where it went wrong.

What is the error?

Computers and Language
Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA
Language learning
Error analysis
Feedback
Student model

CALL systems
Basic uses of computers
Early CALL systems

ICALL
NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

Identifying the exact error can be difficult

- ▶ Take this hypothetical example of someone learning English:
 - (2) Teacher liked to swim.
 - ▶ Did the learner mean:
 - ▶ *Teachers liked to swim.* (problem with plurals)
 - ▶ *The/A teacher liked to swim.* (problem with determiners)
- ⇒ Finding an error is one thing. Figuring out what the learner meant is another (similar to spell checking).

Number of errors

Heift 2001 reports that 40% of sentences for German learners contained more than one error.

- ▶ Don't want to overwhelm students with too much feedback.
- ▶ So, systems can present one error at a time.
 - ▶ Instructors can divide the errors into primary errors and secondary errors – or rank them in some way
 - ▶ Rankings can be based on:
 - ▶ Instructional model, i.e., what is currently being focused on in class
 - ▶ Learner model, i.e., what problems this particular student has
 - ▶ General model, i.e., what problems are common, perhaps for native speakers of a particular language

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

49 / 52

Do we tailor the system to a type of learner?

- ▶ Some systems write mal-rules specifically designed to handle a particular type of learner, e.g. Korean learners of English.
- ▶ Can look at corpora and find the most common errors
 - ▶ Can create an **error typology** = a classification of errors into different groups.
- ▶ Main problem is a lack of generality.

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

50 / 52

Do we tailor the system for an individual?

Do we keep track of a **student model** = what level each student is at, for a given task?

- ▶ Allows us to say which grammatical points need more work for a given individual.
- ▶ Allows us to give different feedback based on the learner's abilities.
- ▶ Make sure the learner knows the terminology presented in the feedback, e.g., for the input *John are big*.
 - ▶ Beginner: "John is a singular subject and are is a plural verb."
 - ▶ Intermediate: "There is a subject-verb agreement error."
 - ▶ Advanced: "There is an error."

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

51 / 52

Constraining the domain

Systems like the e-Tutor work well because they constrain what it is that the students can talk about.

- ▶ Input is fairly free, but students select from a pool of vocab and grammatical structures.
- ▶ Students are fairly introductory, so no exercise is too complex.
 - ▶ Compare teaching people subject-verb agreement vs. teaching them counterfactual clauses in English.
- ▶ All the pre-processing steps (before the grammar check) allow the grammar checker to assume well-formed input.

Computers and Language

Topic 7: CALL

What is ICALL?
Awareness of language forms
An opportunity for CALL

SLA

Language learning
Error analysis
Feedback
Student model

CALL systems

Basic uses of computers
Early CALL systems

ICALL

NLP uses for CALL
Authentic Text ICALL
Parser-Based ICALL

Parser Issues

52 / 52