Language and Computers (Ling 384)

Topic 6: Computer-Aided Language Learning

Adriane Boyd*

Department of Linguistics, OSU

Autumn 2005

* The course was created by Markus Dickinson, Detmar Meurers and Chris Brew.

Outline

Language learning

In order to talk about using computers for language learning, we must figure out what we are talking about:

- First language acquisition (L1)
- **Second language acquisition** (L2)

First language acquisition

How do babies learn language?

- Lack of explicit instruction?
  Positive evidence of language: children learn language based on all the possible sentence around them.
- Parents generally correct content, not grammar.
- Motherese (“baby talk”): different lexicon, intonation, topics, turn-taking
- Tune in on relevant distinctions: synapses still taking shape in the brain

Stages of learning

Babies typically follow the same general stages of learning

- babbling at 6 months
- first words around 11-12 months (able to comprehend more than able to produce)
- two-word stage around 18-24 months
- utterances of varied length, vocab increasing, etc.

Second Language Learning

Second language learning differs from first language acquisition:

- explicit knowledge of a language
- conscious of learning process
- formal teaching helps

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 Langauge 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 learner stage important for developing CALL systems.

Language-specific stages of learning

- 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*

Differences between L1 and L2 acquisition

- We have already mentioned **implicit** vs. **explicit instruction**.
- The best balance of the two for L2 learners is still being debated.
- The two extremes are:
  - “Drill-and-kill”: repetition of exercises
  - Immersion: all the learner needs is well-formed input
- Adult learners are susceptible to **transfer errors** = something from L1 interfering with L2.
  - e.g., East Asian speakers typically mix up *a* and *the* in English: no such distinction in their language.

What is CALL good for?

- Where does computer-aided language learning (CALL) fit in?
- Wherever foreign language teaching is unavailable, inconvenient, or unaffordable.
- CALL can be used in different setups, in particular:
  - self-study
  - supplement to in-class learning
- CALL is a big business: 106 million Euro (about $120 million) spent on CALL products in Europe in 1994. US market is twice as big. (Nerbonne 2003)

What are our expectations?

- Superior to traditional methods of language learning and teaching?
  "Conversation practice machine" (Atwell 1999)
- Supplement to traditional methods?

Using computers to help learn language

Can use:

- 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

Types of exercises

- You can try to program any kind of exercise used in regular foreign language teaching, e.g.:
  - Given the infinitive, use the verb form in a sentence.
  - Point out the errors in this sentence.
  - Write an essay. (More difficult to correct.)
- General guideline: Best to focus on topics covered in class.
- Exercise types (with automated feedback) are limited by how sophisticated your system is.
- We'll return to this issue once we've introduced different kinds of CALL systems.

Frame-based systems

- **Frame-based systems** "match student answers with a set of correct and incorrect answers stored in a frame"
- One can distinguish several types:
  - linear systems
  - branching systems
  - generative systems
- Typical for early CALL systems.

Linear systems

1. pose a question
2. accept an answer
3. inform 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

Essentially layers of linear systems. Student enters:

- Correct response: stay on the same layer → ask the next question in that linear system.
- Incorrect response → system jumps (or branches) to the question in the layer below to which the current question is linked.
  - If the question in the lower level is answered correctly, the system returns to the higher level.
  - If the lower-level question was answered wrongly, a jump is made to a yet lower level offering.

⇒ Much more arduous to come up with the question sets.

Generative systems

- Generate new questions each time system is used. Usually don't have the same session twice
- Based on some **algorithm** = sequence of commands, in this case used to generate new questions

Problems with frame-based systems

- No deep understanding of question domain.
- Merely match answers with questions.
  ⇒ Could be more than one right answer, as with a translation task.

What is a correct answer to an exercise?

Take this fill-in-the-blank exercise (Heift):

- Today is November 10. What date is tomorrow? Tomorrow is ___.
  - The eleventh. [Correct]
  - November 11. [Incorrect?]

⇒ Have to allow for multiple right answers.

What is a correct answer to an exercise? (cont.)

- Erwin arbeitet in Leipzig, aber seine Familie wohnt in Bad Harzburg. Am Wochenende fährt er nach Hause. Erwin fahrt mit dem ___. ?
  (Erwin works in Leipzig but his family lives in Bad Harzburg. On the weekend he drives back home. Erwin takes his ___. ?)
  - Auto (car) [Correct]
  - Wagen (car) [Incorrect?]

⇒ The definition of what is correct is again too rigid.

Intelligent Computer-Aided Language Learning (ICALL)

- Intelligent CALL focuses on using linguistics and natural language processing to make CALL better.
- Examples include:
  - Concordancers
  - Text alignment
  - Speech recognition and synthesis
  - Morphological processing
  - Syntactic processing

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

Text alignment

Show learners texts which are aligned between two languages.

- Advanced learners might benefit by seeing how word usage in their native language correspond to word usage in L2.
- Beginning learners would be overwhelmed.

Speech recognition and synthesis

- 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.

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)

Syntactic analysis

- **Syntactic generation** = system creates sentences based on lemmas/words
  ⇒ create exercise material
- Syntactic parsing
  ⇒ clarify linguistic structure
  ⇒ spot and diagnose errors in learner input

Syntactic processing allows us to deal with unrestricted learner input, such as essays, and give sensible feedback for errors.
Morphological and syntactic processing can also help make the learner more aware of what language is made up of. → **meta-linguistic knowledge** = knowledge about language.

Problems of Syntactic Processing

Main problem with syntactic processing: too many analyses.
I saw the mouse in the house by the garden.
→ *by the garden* can modify *saw*, *mouse*, or *house*.
When learners type in incorrect sentences, you may have to allow for even more analyses.

Error analysis

- So, we have a lot of different technology we can use.
- For many of the exercises we will want to use, the user types in something, and, using some technology, we want to find the error(s) in it.
- Two main issues:
  - Error recognition: What is the error?
  - Feedback: What do you do about the error? (What do you tell the learner?)

Types of systems for error recognition (and feedback)

Heift (Intelligent Language Tutoring Systems for Grammar Practice)

- systems which present only the correct answer → no attempt to find an error
- systems which pinpoint the error by a letter-by-letter comparison between student's answer and correct answer (linear systems)
- systems which anticipate wrong answers (mal-rules)
- systems which use NLP and provide linguistic analysis of sentence

Error recognition issues

- 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?

How do we adapt our technologies to find errors?

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.
- 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 the linguistic reason why not.

Do we tailor the system to a particular kind 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.

Do we tailor the system for an individual learner?

i.e. 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. *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."

What is the error?

- Take this hypothetical example of someone learning English:
  Swimmer liked to swim.
- Did the learner mean:
  - *Swimmers liked to swim.* (problem with plurals)
  - *The/A swimmer 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.
- Can present one error at a time. Instructors can divide the errors into primary errors and secondary errors – or rank them somehow.

An example system

Example system from Heift and Nicholson (2001), which is general (any native language) and which is able to capture different kinds of errors . . . because the exercises are very constrained (as we will see later)

Student Input →

- String match: if the input matches a pre-defined correct answer, we know it's good.
  ⇒ Prevents time-consuming analysis for perfect answers.
- Punctuation check

More on system architecture

- Spell check: run an off-the-shelf spell checker on the input and get the **lemmas** = baseforms of words for the next step.
  - 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.
- Example check
- Missing word check
- Extra word check
  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.

More on system architecture (cont.)

- Word order check: match the user word order with the correct word order (a big issue in German)
- 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.
- 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 below.
- Modularity of the approach.

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)
  Note that negative reinforcement can be bad: “WRONG!”
- Learning processes need feedback to know right from wrong (cognitivism)

There are differences between human and computer feedback

- humans aware of exact student situation.
- humans can infer intentions.

Feedback (cont.)

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.
- Explanations should be short.

Kinds of feedback

- **Explicit correction** = explicitly giving the correct form, indicating that this is a correction.
  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 halba español.
  TEACHER: Yo hablo español tambien. (I speak Spanish also.)
- **Clarification request** = asking for a clarification.
  What? What did you mean?

Kinds of feedback (cont.)

- **Metalinguistic feedback** = comments, information, or questions about the correctness of the student’s utterance, but not giving the correct form.
  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.
  - Yo habl**a**?

Kinds of exercises

Here are some example exercises from a German system (Heift), outlined in
http://www.spz.tu-darmstadt.de/projekt_ejournal/jg_06_2/beitrag/heift2.htm

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

Dictation

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

Dictation (cont.)

Good points:

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

Bad points:

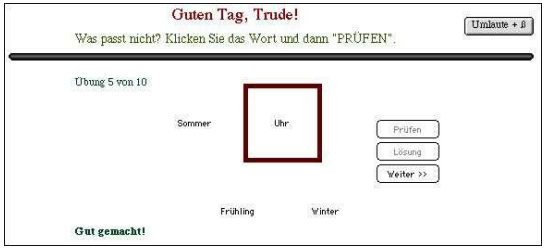
- Requires multimedia resources.
- Takes a long time to prepare.

Build a Phrase

Build up a complete phrase (e.g. NP, but not a whole sentence) based on a given picture (i.e. provide your own vocab).

Which Word is Different

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



Word Order Practice

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

Fill-in-the-Blank

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

Build a Sentence

Use all the given words (lemmas) and create a grammatical German sentence.
Advanced learner output ("There is an error in gender with the subject."):

Build a Sentence (cont.)

Beginning learner output ("No, DER of DER ZEIT is incorrect. ZEIT is not masculine"):

Constraining the Domain

As we said before, Heift's system works pretty well because she constrains 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.

Demos

- Tools:
  - Concordancer
  - Morphological Analyzers XEROX online Demo (English)
  - Parser
- ICALL demos:
  - GLOSSER