Some ICALL example systems

L485/L700

Dept. of Linguistics, Indiana University

Autumn 2008

Example ICALL systems

Can divide them roughly into 2 components:

- Those which elicit learner productions
- Those which have learners interact with authentic text

Since we are focusing on analyzing learner language, we will focus on exercise-based systems

Onilne workbooks

Online learner workbooks typically work as follows:

- System presents the learner with an exercise
- Learner inputs a potentially ill-formed answer
- The parser processes this sentence & identifies problems
- Feedback is then presented to the student

We'll look at 3-4 different example systems:

- e-Tutor (German Tutor): Heift and Nicholson (2001) (German)
- BANZAI/Robo-Sensei: Nagata (2002) (Japanese)
- TAGARELA: Amaral and Meurers (2006, 2007) (Portuguese)
  - Boltun: Dickinson and Herring (2008) (Russian)

e-Tutor (German Tutor)
Heift and Nicholson (2001)

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

- Generality: allows for different native languages (L1)
  - Although most students have L1 English, feedback is in German
- Adaptability: feedback changes based on learner strengths and weaknesses
  - e.g., Consistent subject-verb accuracy leads to more advanced feedback
- Intelligence: HPSG-style parsing system captures errors via the use of mal-rules
  - Able to capture different kinds of errors because the exercises are very constrained
- Persistence of learner modeling: track acquisition patterns, compare longitudinal data
- Modularity: feedback module, e.g., kept separate from answer processing component

Dictation

Some example exercises from e-Tutor (http://zif.spz.tu-darmstadt.de/jg-06-2/beitrag/heift2.htm)

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

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.

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 here: “There is an error in gender with the subject.”

General points

- Provides some contextualized practice outside of a classroom setting (e.g., listening)
- In all of the exercises, processing can be very constrained, leading to tractable processing
  - e.g., listening is limited to those vocabulary words
  - A question to ask is: how about less constrained exercises?
- Feedback is tailored to the learner level
  - The mechanism for this is to assign a learner score for each student and increment/decrement based on performance, for a given construction
  - The score determines a learner's level: 0-10 = beginner, 11-20 = intermediate, 21-30 = advanced

System architecture

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?

System architecture (cont.)

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.

System architecture (cont.)

7. Word order check: match the user word order with the correct word order
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

The error grammar

A quick note on how the grammar works:

- It is a head-driven phrase structure grammar (HPSG) formalism
- The feature architecture is rewritten to store where errors are located
  - [NUM *sg*] becomes [NUM|SG *correct*] or [NUM|SG *error*]
  - A separate feature path is added which keeps track of any errors
- Essentially, this is a way to write mal-rules

Questions for general learner language analysis

Modularity is key if we are to "plug in" new methods for analyzing learner language

- Can new (general) error diagnosis systems be easily integrated into an ICALL system?
- Does the ICALL system have a way of capturing probabilistic judgments of acceptability?
- How can the technology be made more accurate given domain constraints?

BANZAI/Robo-Sensei
Nagata (2002)

Robo-Sensei (formerly, BANZAI) is proprietary software developed for teaching introductory Japanese courses

- Robo-Sensei allows for some multimedia input/output and input/output of Japanese characters
- Lessons are unified by different cultural themes (e.g., department stores)
  - Instructions are given in English, as well as grammatical explanations, presented before any exercises
- Exercises cover a range of Japanese constructions
  - Learners can click on :
    - *Vocabulary* for vocabulary help
    - *Grammar* to re-read the lesson explanation
    - *Feedback* to receive feedback
  - Extensive metalinguistic feedback is given to the learner
    - Some research that more difficult constructions (e.g., particles) are better served with this kind of feedback

Exercises

Sentence-level exercises are the ones which require NLP analysis

- While production is basically free, it is unclear how free, as exercise examples sometimes to be pseudo-translation tasks

> *1. Respond to the following conversation, according to the situation described. In this exercise, do not drop any phrase particle.*
> *(1) Your friend has asked if you ate lunch. Tell him that you ate sushi in Japan Town.*
> *(Nagata 1995)*

Robo-Sensei processing

- 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
- Steps involved in analysis:
  - Word segmenter
  - Error detector: unknown/missing/unexpected words, predicate conjugation errors
  - Parser, applying each phrase structure rule consecutively
  - Error detector

Example grammatical rule

Nagata (1995)

```
<S> <==> (<NP><S>)
  ((((x1 place) =c +)
    ((x1 particle) =c ni)
    ((x1 syn feedback) > (x1 value))
    ((x2 cat) =c verb)
    (*EOR* (((x2 place-ni) =c +))
        ((x1 syn feedback) > this-is-location-ni))
      (((x1 syn feedback) > (x1 value))
        ((x1 syn feedback) > should-be-de-not-ni)))
    ((x0 place) = x1))
  ...
  (x0 = x2)))
```

Some questions

- How can we make processing more general, while at the same time capturing the distinctions that are needed, e.g., for particle usage?
- How much of a target answer do we need information about? Can we get by without it?

TAGARELA

Amaral and Meurers (2006, 2007)

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

- Motivated by providing grammar lessons outside of class, so as not to "reduce the pace" of an in-class communicative lesson
- Instructors wanted tools which:
  - allow for practicing of receptive skills
  - reinforce acquisition of forms
  - raise linguistic awareness in general
- It features standard exercises, as found in foreign language workbooks

Constraining learner input

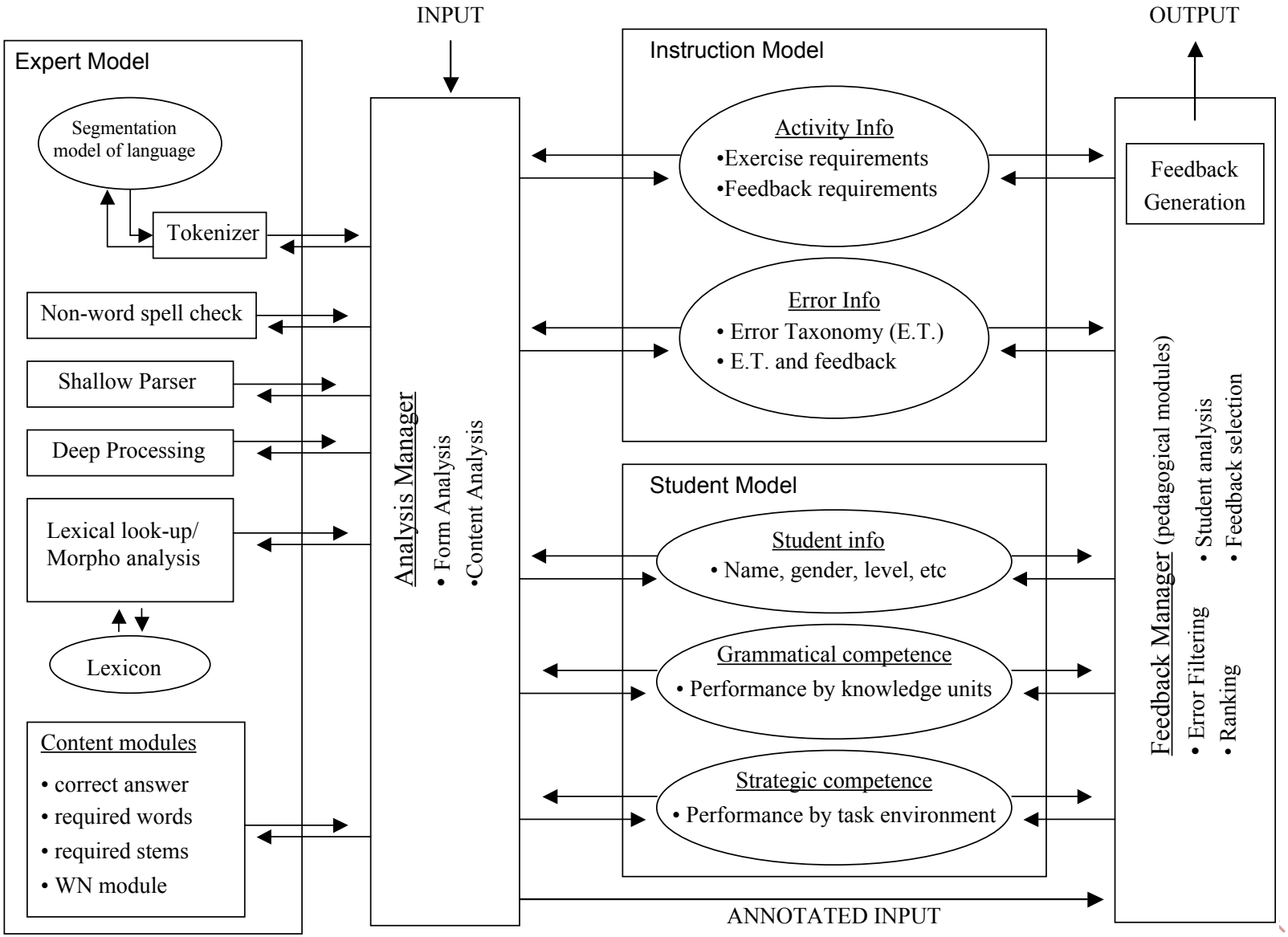
Competing goals (see also Dickinson et al. to appear):

- Provide contextualized exercises which allow for variability in learner input
- Provide accurate & intelligent feedback for any type of error

Goals of TAGARELA:

- Be pedagogically sound in controlling the input
  - Exercises are communicatively significant
  - Communicative setup of the activity (e.g., pictures, instructions) constrains form & content of response
- Explicit learner modeling also points to the types of errors the system needs to be able to deal with

TAGARELA system overview



Demand-driven architecture

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

Linguistic modules

- Form analysis: tokenizer, lexical/morphological lookup, disambiguator, parser
  - NLP modules can be flexibly employed, e.g., interleaved
  - Sometimes multiple analyses from one layer of linguistic structure are preferred
  - In Portuguese, *a* is POS ambiguous (preposition, pronoun, article, abbreviation), but tokenization can resolve some of this (e.g., $da = de + a$)
- Content analysis: shallow semantic matching

Activity models

Activities are characterized in terms of:

- task specification
- level
- expected input
- nature & availability of target reponses & type of permitted variation
- required skills & abilities
- pedagogical goals

Depending upon the activity, different kinds of linguistic processing are needed and will be more or less accurate

Annotation-based processing

NLP analysis = a process of enriching the learner input with annotations

- Which annotation module needs to be called when is determined by a controller using the input & output specifications of each module
- Error analysis is done separately from linguistic annotation

Note that shift in emphasis:

- Broad-coverage NLP tools can be employed when they do not always have to guess at what is right or wrong about a sentence.
- The key issue is whether such tools provide relevant information and information which supports analyzing learner language

Questions

- What types of tools readily support providing annotation to learner language?
- When are robust tools beneficial? When will they be harmful in analyzing learner productions?
- What is the line between error analysis and linguistic analysis?
- Should our emphasis be on detecting/diagnosing errors or only on supporting such endeavors? How do we do that?

References

Amaral, Luiz and Detmar Meurers (2006). Where does ICALL Fit into Foreign Language Teaching? CALICO Conference. May 19, 2006. University of Hawaii http://www.ling.ohio-state.edu/icall/handouts/calico06-amaral-meurers.pdf.

Amaral, Luiz and Detmar Meurers (2007). Putting activity models in the driver's seat: Towards a demand-driven NLP architecture for ICALL. EUROCALL. September 7, 2007. University of Ulster, Coleraine Campus. http://www.ling.ohio-state.edu/icall/handouts/eurocall07-amaral-meurers.pdf.

Dickinson, Markus, Soojeong Eom, Yunkyoung Kang, Chong Min Lee and Rebecca Sachs (to appear). A Balancing Act: How can intelligent computer-generated feedback be provided in learner-to-learner interactions. *Computer Assisted Language Learning* .

Dickinson, Markus and Joshua Herring (2008). Developing Online ICALL Exercises for Russian. In *The 3rd Workshop on Innovative Use of NLP for Building Educational Applications (ACL08-NLP-Education)*. Columbus, OH, pp. 1–9. http://jones.ling.indiana.edu/~mdickinson/papers/dickinson-herring08.html.

Heift, Trude and Devlan Nicholson (2001). Web delivery of adaptive and interactive language tutoring. *International Journal of Artificial Intelligence in Education* 12(4), 310–325.

Nagata, Noriko (1995). An Effective Application of Natural Language Processing in Second Language Instruction. *CALICO Journal* 13(1), 47–67. http://www.usfca.edu/japanese/CALICO95.pdf.

Nagata, Noriko (2002). BANZAI: An Application of Natural Language Processingto Web based Language Learning. *CALICO Journal* 19(3), 583–599. http://www.usfca.edu/japanese/CALICO02.pdf.