



NLP & CALL

Advice on NLP in CALL

John Nerbonne, University of Groningen
Computational and Pedagogical Challenges
for NLP in CALL
EUROCALL 2001 Nijmegen
29 Aug. 2001

RUG



This talk

NLP & CALL

- perspective on CALL — 2nd lg. learning
- theory and application, esp. in NLP
- sufficiently reliable technology
- error recognition and diagnosis
- opportunities for error recognition
- need for evaluation

Thesis: parsing is insufficiently reliable for error recognition and diagnosis *in general*



Perspective: CALL is (2nd-)Language Learning^{NLP & CALL}

pace soul-searching à la Levy: Must CALL be viewed as a kind of AI, Cognitive Science, Linguistics, ... ?

(2nd-)Language Learning provides theory, analytical techniques, evaluation standards, application ideas, ...

CALL is simply a technical innovation in (2nd LL, like bilingual dictionaries, learner's grammars, tape recorders, ...



CALL attractive NLP application

NLP & CALL

Language Learning Hard, Worthwhile

- 60 – 100 hours to reach primitive level, (order coffee, get directions)
- improvement costs exponential (in learner's time needed)
- schools, universities, private language schools, corporate & government training, private “edutainment” (> 100 Mil. Euro)
- schools, programs, teachers often unavailable, too expensive

Don't ask “Is CALL better?”, but rather “Where will it succeed?”

Paper sketches an very rough estimate of the market size in multimedia products in CALL, ca. 500 Mil. Euro.

RUG



NLP Applications

NLP & CALL

Applications are relations

technique \Leftrightarrow need
technology market
knowledge

Many “applications” one-sided:

technology with no practical need in sight practical improvement with no leading technical idea



Open Eye

NLP & CALL

Many successful developments are **not** simple automations of existing processes:

sewing machines	unlike tailors' stitching
penicillin	“ruined” culture
airfoils	unlike birds' wings
post-it's	extremely weak adhesive
hypersearch à la Google	unlike library search

Conclusion: CALL needn't mimic teachers, but we need *sufficiently reliable* technology



Language Technology

NLP & CALL

Mode: speech, print, handwriting

Task: recognize, generate, understand, converse, translate, index, correct, search (retrieve), language learning

Medium/Locale: telephone, PC, automobile, toys, factory floor, PDA

Application \in **Mode** \times **Task** \times **Medium**

falling price of hardware, growing demand (Information Highway,...)



Present NLP Applications

NLP & CALL

- **NL Interfaces** Intellect, NAT (Ginsparg, Berkeley), Q&A (Hendrix, Cupertino) Paralance (BBN, Bobrow & Bates, Cambridge)
- **MT SYSTRAN, METAL, LOGOS, BDC EC** (Hsin-Chu, Taiwan)
- **documentation support (“controlled language”)** Boeing (Hoard, Wojcik et al.), Hughes
- **grammar checkers** Hugo+, Correct Grammar, Grammatik 5, GramR, ...
- **Information Retrieval** Whiz Bang! et cetera
- **Automatic Email Answering** YY
- **spelling checkers, hyphenators**
- **handicapped aids**
- **DB address book**
- **OCR correction** (Xerox)

little directly on CALL (but some in Syracuse Language Systems)



Unrealistic Expectations

NLP & CALL

[...] 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

This can lead to frustration, and unwarranted dissatisfaction.



Technically Feasible CALL with NLP

NLP & CALL

- concordancing
 - Lancaster's work (McEnery)
 - Barlow's *Monocore*
- alignment
- speech recognition for word pronunciation (EUROCALL 2000)
- morphological analysis

concordancing is straightforward, and solved



GLOSSER — Reading Assistance

NLP & CALL

- for intermediate-level Dutch learners of French
- arbitrary French texts
- assistance
 - dictionary entry French-Dutch *Van Dale*
 - gram. significance of inflection
 - examples of word use from corpora
- reduced functionality in Web Demo
www.let.rug.nl/~glosser/



File Edit View Go Communicator Help

Bookmarks Location: <http://odur.let.rug.nl/~gloss/> What's Related

Home Help Morphology Dictionary Examples
Text index On Off On Off On Off

Le chêne et le roseau

Le chêne un jour dit au roseau:
"Vous avez bien sujet d'accuser la nature:
Un roitelet pour vous est un pesant fardeau.
Le moindre vent qui d'aventure
Fait rider la face de l'eau
Vous oblige à baisser la tête;
Cependant que mon front, au Caucase pareil,
Non content d'arrêter les rayons du soleil,
Brave l'effort de la tempête.
Tout vous est aquilon, tout me semble zéphyr.
Encor si vous naissiez à l'abri du feuillage
Dont je couvre le voisinage,
Vous n'auriez pas tant à souffrir:
Je vous défendrais de l'orage.
Mais vous naissez le plus souvent
Sur les humides bords des royaumes du vent.

Analysis:
naître+Imp+PL +P2+FinV

naître [nɛ : tr] t60 0.1 geboren
worden => ter wereld komen 0.2 fig.
ontstaan => geboren worden,
opkomen 1.2 *faire* ~ *une industrie* een
industrie oprichten 1.2 *le jour*
commence à ~ de dag breekt aan 6.1
en naissant bij de geboorte .1 *innocent*
comme un enfant qui vient de ~

... absent tout entre, tout se mêle :
Les livres sur Evrard fondent
comme la grêle Qui, dans un grand
jardin, à coups impétueux, Abat
l'honneur **naissant** des rameaux
fructueux. Chacun s'arme au hasard
du livre qu'il rencontre : L'un tient
l'Edit d'amour, l'autre en saisit la

100%



GLOSSER

NLP & CALL

- manageable task, successful implementation
 - technical test on 500 random words
 - * 99% coverage (morphology)
 - * 3% error in disambiguation – choosing possible, but unintended analysis
 - user test showed improved speed
- lukewarm reception from language teachers
 - “can’t teach vocabulary”
 - too grammar oriented
 - unnatural task
 - too easy for learners
- idea incorporated into Syracuse Language Systems
“Learn Spanish Your Way”



Comprehensible Input

NLP & CALL

Krashen's "Fundamental Pedagogical Principle":

foreign language input must be comprehensible for learning to occur

Krashen goes on to speculate that this may be (nearly) sufficient — since automatic processes of “acquisition” then apply

Critics argue about this, but not about **necessary** condition.

role for language technology: facilitate comprehension



Recognizing and Correcting Errors

NLP & CALL

important to language pedagogy

Allen “Ten Desiderata for CALL” *Computers and the Humanities* 96/97

1. unimportant errors should be corrected and then ignored
2. programs should allow variant answers
3. programs should analyze students’ answers as they’re written.
7. programs should comment on common errors, not just errors w.r.t. a particular question
8. programs should be sensitive to syntax in order to facilitate error recognition
9. programs should ignore typing errors



Recognizing and Correcting Errors

NLP & CALL

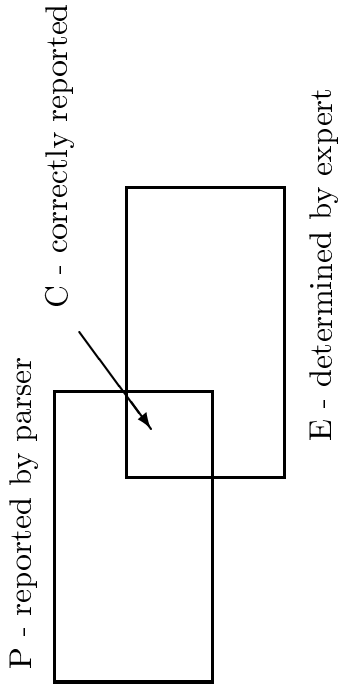
general error diagnosis and correction is hard

- syntactic analysis is hard
- massive disambiguation problem
- added complexity of errors

note disappointments with “grammar checkers” (Tschichold '99)



Syntactic Analysis



Correctly found: $C = P \cap E$

Precision = C/P ; Recall = C/E

Average Precision and Recall reported ($F_{\beta=1}$)

But what is reported: Sentences, Constituents, Messages?



Performance

NLP & CALL

- unrestricted correct written input (WSJ) $F = 90\%$
 - evaluations report on **constituents**, not sentences — $F = 45 - 50\%$ on sentences
- very restricted natural spoken input (telephone information system) $\approx 85\%$
 - note that this includes speech recognition (92% of what's possible)
 - “From what city to what city would you like to travel?”
 - evaluation report on messages (pragmatic guessing OK)

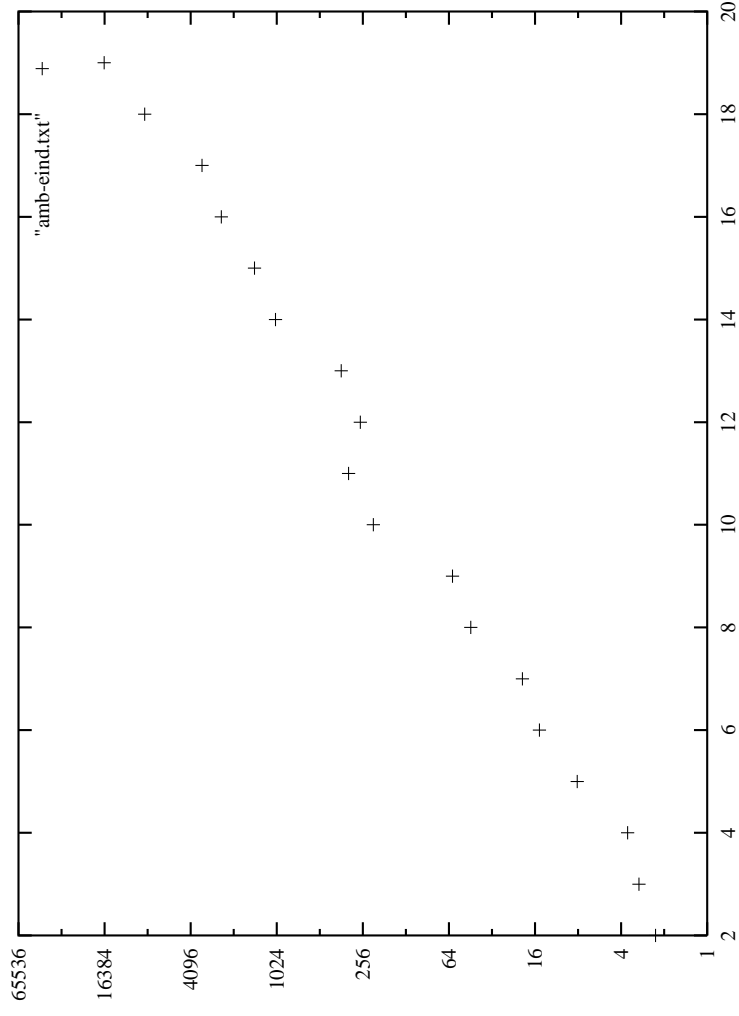
Conclusion: low level of accuracy for natural, even edited syntax.



Disambiguation

NLP & CALL

Average number of parses as function of sentence length (loglinear).





Disambiguation

NLP & CALL

- Dutch newspaper text (Eindhoven corpus)
- Alpino Grammar (van Noord, Bouma, Malouf)
 - best Dutch grammar in competition

Possible parses increase exponentially as a function of sentence length. Even 10-word sentences are significantly ambiguous.

Very careful syntactic analyses underdetermine intended structure. Successful systems disambiguate statistically, but require training material.

Conclusion: Attend to parser evaluation wrt disambiguation.



Techniques for Error Recognition

NLP & CALL

- *mal*-rules to cover ill-formed structures
- constraint relaxation

What do you see? I see **a/the** bird.
 $S \rightarrow (NP) VP$
 $NP \rightarrow (Det) N$

note imperative, NN compound analyses



Disambiguation and Ill-Formed Input

NLP & CALL

Allowing for errors **increases** the number of parses.

Mary leave tomorrow.

Mary leaves tomorrow.

Mary, leave tomorrow. (imperative)

Some real text (from a colleague's web site):

Lastly you get a link to the **weeks** assignment. [...]

To do so there are two **courses** every week (\emptyset some exceptions for the practical). In the lecture the more theoretical backgrounds will be given, \emptyset this theory will be applied in the practical where you can do the trick yourself. Attending **to** both the lecture and practical is compulsory, \emptyset your presence **to** the course will determine 20% of the final grade. If you can not come to a class **with** a good reason (illness etc.) \emptyset **it** will not be calculated.



Coping with Massive Ambiguity in NLP

NLP & CALL

- reduce coverage (“sublanguage”)
 - Levin & Evans’s ALICE-Chan
 - needs careful, special development
- use statistical training
 - requires annotated training material
 - no application in CALL to-date

systematic evaluation is lacking or limited (in CALL applications)



Premature Rejection

NLP & CALL

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.

Salaberry, 1996, p.12



Practical Alternatives (Ideas)

NLP & CALL

- reduce scope
 - error recognition in exercises whose answers are single words, or short phrases
 - no unrestricted input
- design exercises so that error-checking is facilitated
 - Express in perfect tense.

systematic evaluation is lacking or limited (in CALL applications)



Premature Rejection

NLP & CALL