

# Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amaral and Detmar Meurers

The Ohio State University

EUROCALL  
September 6, 2006

Using Foreign  
Language Tutoring  
Systems for  
Grammatical  
Feedback

Luiz Amaral & Detmar Meurers

Introduction

TAGARELA  
Pedagogical  
Components

Electronic Workbook  
Exercise Types

Processing  
Student Input

Why use NLP?  
Basic Characteristics of ITS

TAGARELA  
System  
Components

System Architecture  
Expert Model  
Student Model  
Instruction Model  
Feedback Manager

Previous Work

Our Contribution

References



1 / 23

## Introduction

- ▶ The Ohio State University has the longest running individualized language instruction (I.I.) center in the US.
- ▶ In 2004 Portuguese was added as a language taught.
- ▶ It provided a good opportunity to investigate the inclusion of an intelligent CALL (ICALL) system.
- ▶ In 2005 we began the implementation of TAGARELA, an ICALL system to be integrated into I.I.
- ▶ This presentation discusses the necessary components of an ICALL system to fulfill its pedagogical goals.

Using Foreign  
Language Tutoring  
Systems for  
Grammatical  
Feedback

Luiz Amaral & Detmar Meurers

Introduction

TAGARELA  
Pedagogical  
Components

Electronic Workbook  
Exercise Types

Processing  
Student Input

Why use NLP?  
Basic Characteristics of ITS

TAGARELA  
System  
Components

System Architecture  
Expert Model  
Student Model  
Instruction Model  
Feedback Manager

Previous Work

Our Contribution

References



2 / 23

## TAGARELA in I.I. Program

### *Teaching Aid for Grammatical Awareness, Recognition and Enhancement of Linguistic Abilities (TAGARELA)*

- ▶ TAGARELA uses natural language processing (NLP) technology to process students' input.
- ▶ We have designed TAGARELA to fill some of the pedagogical shortcomings of the I.I. setup.
  - ▶ TAGARELA offers on the spot individualized feedback on spelling, morphological, syntactic and semantic errors.
  - ▶ It provides opportunities for students to practice their listening, reading, and writing skills.

Using Foreign  
Language Tutoring  
Systems for  
Grammatical  
Feedback

Luiz Amaral & Detmar Meurers

Introduction

TAGARELA  
Pedagogical  
Components

Electronic Workbook  
Exercise Types

Processing  
Student Input

Why use NLP?  
Basic Characteristics of ITS

TAGARELA  
System  
Components

System Architecture  
Expert Model  
Student Model  
Instruction Model  
Feedback Manager

Previous Work

Our Contribution

References



3 / 23

## How does TAGARELA fit into I.I.?

- ▶ TAGARELA can be viewed as an intelligent electronic workbook
- ▶ Its exercise types are similar to the ones found in regular workbooks:
  - ▶ Listening
  - ▶ Reading
  - ▶ Description
  - ▶ Fill in the Blanks
  - ▶ Rephrasing
  - ▶ Vocabulary

Using Foreign  
Language Tutoring  
Systems for  
Grammatical  
Feedback

Luiz Amaral & Detmar Meurers

Introduction

TAGARELA  
Pedagogical  
Components

Electronic Workbook  
Exercise Types

Processing  
Student Input

Why use NLP?  
Basic Characteristics of ITS

TAGARELA  
System  
Components

System Architecture  
Expert Model  
Student Model  
Instruction Model  
Feedback Manager

Previous Work

Our Contribution

References



4 / 23

THE TAGARELA SYSTEM @ THE OHIO STATE UNIVERSITY ICALL RESEARCH GROUP

Listening Reading Description Fill-In-Blanks Rephrasing Vocabulary Home Logout


## Leitura

Módulos: 1 2 3 4 5 Atividades: 1 2

Instrução

Leia o texto e responda às questões usando frases completas e o vocabulário apresentado no texto.

English Instructions



**Regiões do Brasil**

O Brasil está política e geograficamente dividido em cinco regiões. Os limites de cada região (Norte, Nordeste, Sudeste, Sul e Centro-Sul) coincidem sempre com os limites dos estados que os compõem.

A região Norte ocupa a maior parte do território brasileiro, com uma área que corresponde a 45,27% da área total do País. Formada por sete Estados, tem sua área quase totalmente dominada pela bacia do Rio Amazonas.

A região Nordeste pode ser considerada a mais heterogênea do País. Dividida em quatro grandes áreas - meio-norte, zona da mata, agreste e sertão - ocupa 18,26% do território nacional e tem nome antigo.

O Sudeste é formado por quatro Estados. Esta é a região de maior importância econômica do País, onde está concentrado também o maior índice populacional - 42,8% dos brasileiros.

Já o Sul, região mais fria do País, com ocorrência de geadas e neve, é a que apresenta menor área, ocupando 6,7% do território brasileiro e com apenas três Estados. Os rios que cortam sua área formam a bacia do Paraná em quase toda sua extensão e são de grande importância para o País, principalmente pelo seu potencial hidroelétrico.

Finalmente, a região Centro-Oeste tem sua área delimitada basicamente pelo Planalto Central Brasileiro e pode ser

Questão 1

Quantas regiões tem o Brasil?

Resposta:

Enviar

Problema Questão (0)

Análise:

Ajuda Dicionário Gramática

THE OHIO STATE UNIVERSITY

8 / 23

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amarel & Detmar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why Use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

THE OHIO STATE UNIVERSITY

5 / 23

THE TAGARELA SYSTEM @ THE OHIO STATE UNIVERSITY ICALL RESEARCH GROUP

Listening Reading Description Fill-In-Blanks Rephrasing Vocabulary Home Logout

## Compreensão Auditiva

Módulos: 1 2 3 4 5 Atividades: 1 2

Instrução

Ouçe o diálogo e complete a ficha com o que cada cliente pediu.

English Instructions



Questão 1

BEBIDA:

ENTRADA:

PRATO PRINCIPAL:

GUARNIÇÃO:

SOBREMESA:

Enviar

Questões: 1

Análise:

Ajuda Dicionário Gramática

THE OHIO STATE UNIVERSITY

6 / 23

THE TAGARELA SYSTEM @ THE OHIO STATE UNIVERSITY ICALL RESEARCH GROUP

Listening Reading Description Fill-In-Blanks Rephrasing Vocabulary Home Logout

## Descrição

Módulos: 1 2 3 4 5 Atividades: 1 2

Instrução

Descreva a foto usando as palavras apresentadas no exercício e uma das preposições abaixo.

em cima de - entre - embaixo de - ao lado de

English Instructions

Questão 1



(vaso - mesa)

Resposta:

Enviar

Problema Questão (0)

Análise:

Ajuda Dicionário Gramática

THE OHIO STATE UNIVERSITY

7 / 23

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amarel & Detmar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why Use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

THE OHIO STATE UNIVERSITY

8 / 23

THE TAGARELA SYSTEM @ THE OHIO STATE UNIVERSITY ICALL RESEARCH GROUP

Listening Reading Description Fill-In-Blanks Rephrasing Vocabulary Home Logout

## Preencha as Lacunas

Módulos: 1 2 3 4 5 Atividades: 1 2

Instrução

Complete as lacunas com os verbos listados abaixo. Não repita o mesmo verbo mais de uma vez. Conjugue os verbos no pretérito perfeito do Indicativo.

English Instructions

Questão 2

fazer - estar - gostar

Eu já  em Nova York duas vezes. Eu  uma viagem em 1998 e outra em 2002.

Eu  muito da cidade.

Enviar

Problema Questão (0)

Análise:

Ajuda Dicionário Gramática

THE OHIO STATE UNIVERSITY

8 / 23

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amarel & Detmar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why Use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

THE OHIO STATE UNIVERSITY

9 / 23

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amarel & Detmar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why Use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

THE OHIO STATE UNIVERSITY

10 / 23

THE TAGARELA SYSTEM @ THE OHIO STATE UNIVERSITY CALL RESEARCH GROUP

Listening Reading Description Fill-In-Blanks Rephrasing Vocabulary Home Logout

**Reescrava**

Módulos: **1 2 3 4 5** Atividades: **1 2**

**Instrução**

Escreva uma frase comparando os dois elementos apresentados na tabela. Siga o exemplo abaixo.

ensolarada
cozinha 1: sol todo o dia
cozinha 2: sol só pela manhã

Resposta: A cozinha 1 é mais ensolarada que a cozinha 2.

**Questão 1** Questões: **1 2 3 4 5 6 7 8** Análise:

caro	Ajuda Dicionário Gramática
apartamento 1: R\$150.000	
apartamento 2: R\$230.000	

Enviar

Próxima Questão (2)

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amaral & Dámar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

OHIO STATE

9 / 23

THE TAGARELA SYSTEM @ THE OHIO STATE UNIVERSITY CALL RESEARCH GROUP

Listening Reading Description Fill-In-Blanks Rephrasing Vocabulary Home Logout

**Vocabulário**

Módulos: **1 2 3 4 5** Atividades: **1 2 3**

**Instrução**

Observe a figura e complete a descrição com as palavras que estão faltando.

English Instructions

**Questão 1** Questões: **1 2 3 4 5 6 7 8 9**



No banheiro tem  Enviar

Próxima Questão (2)

**Análise:**

Ajuda Dicionário Gramática

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amaral & Dámar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

OHIO STATE

10 / 23

## Why use natural language processing?

- ▶ String matching is the most common technique used in CALL to analyze student input, which works well when
    - ▶ correct answers and potential errors are predictable and listable
    - ▶ there is no grammatical variation
    - ▶ envisaged errors correspond directly to intended feedback
  - ▶ But what if
    - ▶ possible correct answers are predictable but not (conveniently) listable for a given activity
    - ▶ errors can occur throughout a recursively built structure
    - ▶ personalized feedback is desired which requires information about the learner input that can only be obtained through linguistic analysis
- Use natural language processing to analyze student input in such cases.

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amaral & Dámar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

OHIO STATE

11 / 23

## Intelligent Tutoring Systems

- ▶ An Intelligent Tutoring System (ITS) is a computer program that intelligently interacts with the learner.
- ▶ An ITS should be able to:
  - ▶ accurately diagnose the knowledge structures and skills of the student
  - ▶ adapt instruction accordingly
  - ▶ provide personalized feedback
- ▶ Since Hartley and Sleeman (1973) an ITS is recognized as consisting of at least three components:
  - ▶ the expert model
  - ▶ the student model
  - ▶ the instruction model

Using Foreign Language Tutoring Systems for Grammatical Feedback

Luiz Amaral & Dámar Meurers

Introduction

TAGARELA Pedagogical Components

Electronic Workbook

Exercise Types

Processing Student Input

Why use NLP?

Basic Characteristics of ITS

TAGARELA System Components

System Architecture

Expert Model

Student Model

Instruction Model

Feedback Manager

Previous Work

Our Contribution

References

OHIO STATE

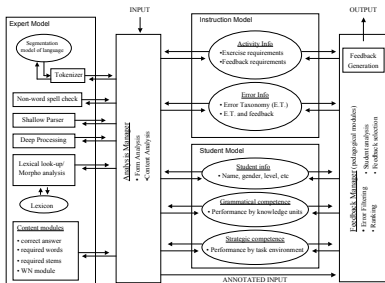
12 / 23

## Components of an ITS

- ▶ **Expert Model:**
  - ▶ the knowledge that the ITS has of its subject domain, in our case the linguistic knowledge
- ▶ **Student Model:**
  - ▶ the component of the system keeping track of student's current state of knowledge
  - ▶ It allows the ITS to infer the student's understanding of the subject matter and to adjust the feedback to the student's needs.
- ▶ **Instruction Model:**
  - ▶ the component that stores pedagogical information, how to conduct instruction
  - ▶ It helps define strategies to deliver appropriate feedback.



## General Architecture of TAGARELA



## Analysis Manager

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



## Expert Model

- ▶ **Tokenizer**
  - ▶ Segmentation model of language
- ▶ **Non-word spell checker**
- ▶ **Shallow Parser**
- ▶ **Deep processing**
- ▶ **Lexical look-up and Morphological analysis**
  - ▶ Lexicon
- ▶ **Content Assessment (Bailey & Meurers 2006)**
  - ▶ correct answer test
  - ▶ required word test
  - ▶ required stem test
  - ▶ WordNet module



## Student Model

- ▶ Personal Information:
  - ▶ name, level, age, gender, etc.
- ▶ Grammatical Competence:
  - ▶ Student's performance on specific knowledge items (parallel to ICICLE, Michaud et al. 2001; E-Tutor, Heift 2004)
- ▶ Strategic Competence:
  - ▶ Student's performance given specific task environment
  - ▶ Task environments defined in terms of:
    - ▶ activity level,
    - ▶ type of input, and
    - ▶ amount of content manipulation required.
  - ▶ New approach in ICALL systems, integrating cognitive modeling proposed by Bachman & Palmer (1996).



## Instruction Model

- ▶ Information about Errors
  - ▶ types of errors
  - ▶ expected level of each error type
  - ▶ expected errors per activity type
  - ▶ best feedback strategies per error type
- ▶ Information about Activities
  - ▶ types of activities
  - ▶ best processing mechanism for each activity type
  - ▶ best feedback strategies for each activity type



## Feedback Manager

- ▶ receives the annotated input from the Analysis Manager
- ▶ gathers the necessary information from:
  - ▶ student model
  - ▶ instruction model
- ▶ decides on the best feedback strategy, generating a feedback log
- ▶ updates the student model
- ▶ composes a feedback message (HTML)



## Previous Work

- ▶ Intelligent Tutoring Systems (cf., e.g., Corbett et al. 1997)
  - ▶ integrate expert, instruction and student models
  - ▶ successful for restricted sub-domains of science
- ▶ CALL systems
  - ▶ focus on language instruction
  - ▶ lack expert model
  - ▶ interaction with student limited to pre- envisaged language
- ▶ ICALL systems (Robo-Sensei, Nagata 2002; E-Tutor, Heift 1998)
  - ▶ focus on expert model
  - ▶ limited incorporation of instructional components
  - ▶ student models based on acquisition of form



## Contributions of our Project

- ▶ The project emphasizes the integration of computational, linguistic, and FLT/SLA expertise in ICALL research.
- ▶ Expert Model:
  - ▶ Demand-driven architecture allows for multiple ways of processing students input depending on the activity type.
  - ▶ Incorporating semantic assessment (Bailey & Meurers 2006)
  - ▶ Advancing constraint-based processing, dealing with a wider range of errors (Metcalfe & Meurers 2006)
- ▶ Student Model:
  - ▶ Integrate broader view of the acquisition process
  - ▶ Modeling strategic competence in addition to grammatical competence
- ▶ Instruction Model:
  - ▶ Explicitly incorporates information about pedagogical aspects of instruction.



## Status of the Project

TAGARELA is a two year project that started in Summer 2005 and will end in Spring 2007. Currently:

- ▶ activity environment is finished
- ▶ exercises are implemented
- ▶ web-based interface is ready to be tested
- ▶ Fall 2006 / Winter 2007:
  - ▶ finish development and testing of NLP components
  - ▶ finish the implementation of learner and expert models
- ▶ Spring 2007:
  - ▶ Integration into the I.I. of Portuguese at OSU
  - ▶ Testing with students enrolled in the program



## Acknowledgements

### We are grateful to

- ▶ the OSU Office of Research and the College of Arts and Humanities for the Innovation Grant funding,
- ▶ the Núcleo Interinstitucional de Linguística Computacional (NILC) for providing access to Portuguese NLP resources, especially Ricardo Hasegawa, Maria das Graças Volpe Nunes and Sandra Maria Alufisio,
- ▶ the Student Technology Consultants Cory Shain and Chris Kovach, and their HIS advisors Scott King, Dickie Selve and Jody Croley for the web design,
- ▶ Glaucia Silva, Donna Long, Terence Odlin, and Alison Mackey for helpful discussion of SLA/FLT, and
- ▶ the participants of the OSU ICALL research group and the OSU CLippers computational linguistics discussion group for helpful feedback on our work.



## References

- Bachman, Lyle F. and Adrian S. Palmer (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford, UK: Oxford University Press.
- Bailey, Stacey and Detmar Meurers (2006). Exercise-driven selection of content matching methodologies. EUROCALL Conference. September 6, 2006. University of Granada.
- Corbett, Albert T., Kenneth R. Koedinger and John R. Anderson (1997). Intelligent Tutoring Systems. In Martin G. Helander, Thomas K. Landauer and Prasad V. Prabh (eds.), *Handbook of Human Computer Interaction*, Amsterdam: North-Holland, pp. 849–874.
- Hartley, J.R. and Derek H. Sleeman (1973). Towards intelligent teaching systems. *International Journal of Man-Machine Studies* 5, 215–236.
- Heift, Gertrud D. (1998). Designed Intelligence: A Language Teacher Model. Ph.D. thesis, Simon Fraser University.
- Heift, Trude (2004). Inspectable Learner Reports for Web-based Language Learning. *ReCALL Journal* 16(2), 416–431.
- Metcalfe, Vanessa and Detmar Meurers (2006). When to Use Deep Processing and When Not To – The Example of Word Order Errors. Pre-conference Workshop on NLP in CALL – Computational and Linguistic Challenges. CALICO 2006. May 17, 2006. University of Hawaii.
- Michaud, Lisa N., Kathleen F. McCoy and Litza A. Stark (2001). Modeling the Acquisition of English: An Intelligent CALL Approach. In *Proceedings of The*



8th International Conference on User Modeling, Sonthofen, Germany, pp. 14–25.

Nagata, Noriko (2002). BANZAI: An Application of Natural Language Processing to Web based Language Learning. *CALICO Journal* 19(3), 583–599.

Using Foreign  
Language Tutoring  
Systems for  
Grammatical  
Feedback

Luis Amoril & Detmar Meurers

Introduction

TAGARELA  
Pedagogical  
Components

Electronic Workbook  
Exercise Types

Processing  
Student Input

Why Use NLP?  
Basic Characteristics of ITS

TAGARELA  
System

Components

System Architecture  
Expert Model  
Student Model  
Instruction Model  
Feedback Manager

Previous Work

Our Contribution

References

