Language Learning Tasks and Automatic Analysis of Learner Language
Connecting FLTL and NLP in the design of ICALL materials supporting effective use in real-life instruction

Martí Quixal (Ph.D. Candidate)

Co-advisors:

Dr. Toni Badia
Departament de Traducció i Ciències del Llenguatge
Universitat Pompeu Fabra

Prof. Dr. Walt Detmar Meurers
Seminar für Sprachwissenschaft
Eberhard-Karls-Universität Tübingen

Dec 20th, 2012
1 Introduction

2 Background

3 ICALL tasks – Where FLTL meets NLP

4 Enabling teachers to author ICALL activities

5 Conclusions
1 Introduction
   - Context and motivation
   - Overview of research in ICALL
   - Goals

2 Background

3 ICALL tasks – Where FLTL meets NLP

4 Enabling teachers to author ICALL activities

5 Conclusions
Figure: Intelligent CALL (ICALL) is at the heart of FLTL, CALL and NLP.
Motivation

- Computer-based language learning can increase motivation and success of language learners

- But, the use of Natural Language Processing in language teaching and learning is disputed and historically controversial
  (Heift and Schulze, 2007: pp. 224–225)

⇒ Compatibility with communicative approaches (FLTL)
⇒ The need for “a good problem to solve” (NLP)
⇒ Lack of interdisciplinary research (ICALL)
  (ten Hacken, 2003; Heift and Schulze, 2007; Amaral and Meurers, 2011)
Weischedel et al. (1978) apply NLP to FLTL by seizing the *natural* restrictions of the “world” described in a text.

Research has shown that:

- It can be integrated in CLT approaches
- Allows for free use of language, and fosters learner autonomy

(Heift and Schulze, 2007; Schulze, 2008, 2010: p. 70–78; Antoniadis et al., 2004; Amaral and Meurers, 2011)
Challenge: *When are ICALL tasks meaningful?*

The middle ground – adapted from (Bailey and Meurers, 2008: p.108)
Overview of research in ICALL

Challenge: *teacher involvement*

- ICALL systems have a limited presence in real-world instruction settings
  - Involve FLTL experts in research and practice (Amaral and Meurers, 2011: p. 19)
  - Lack of authoring tools (Levy, 1997: p. 19)
Research goals

G1  Characterise the pedagogical and computational requirements of successful ICALL activities

G2  Design and evaluate a methodology/technology for teachers to author their own ICALL activities
Introduction

Background
- Natural Language Processing
- Foreign Language Teaching and Learning

ICALL tasks – Where FLTL meets NLP

Enabling teachers to author ICALL activities

Conclusions
NLP approach: domain-adaptive, robust and modular
FLTL approach

- Communicative Language Teaching
  - Task-Based Language Instruction
  - Classification of pedagogical tasks

- Formative and summative assessment of learner production
  - Characterisation of target language use setting
  - Characterisation of language to be elicited from the learner

- Provide corrective feedback
FLTL-driven adaptation of automatic assessment

LEARNER RESPONSE → WORD AND TEXT SEGMENTATION → MORPHOLOGICAL ANALYSER → MORPHOSYNTACTIC DISAMBIGUATION

ASSESSMENT WITH GENERAL NLP RESOURCES

INFORMATION EXTRACTION MODULE(S) → FEEDBACK GENERATION MODULE(S) → ASSESSMENT WITH TASK SPECIFIC NLP RESOURCES

CONTEXT SENSITIVE SPELL CHECKING → SPELL CHECKING

TASK-SPECIFIC AUTOMATIC ASSESSMENT

(TASK = DOMAIN)
Introduction

2 Background

ICALL tasks – Where FLTL meets NLP
- Designing ICALL tasks: Characterisation of pedagogical needs
- NLP functionalities to fulfil FLTL needs
- Learner data to inform and validate design

4 Enabling teachers to author ICALL activities

5 Conclusions
Elements of an ICALL task in the design phase

- **L_T**: Language Task
- **T**: Task
- **ACT**: Activity
- **EXR**: Experience
- **AA**: Action
- **LA**: Language Activity
- **FG**: Feature Group

**PROCESSING CAPABILITIES**

**PEDAGOGICAL GOALS**
**TAF: Task Analysis Framework**

Characterise FL learning activities to:

a) define their communicative and linguistic goals at a broad level

b) pre-select candidates for NLP-based assessment
Designing ICALL tasks: Characterisation of pedagogical needs

RIF: Response Interpretation Framework I

Characterise responses to learning activities in terms of:

a) input data
b) topical and linguistic knowledge
c) pedagogically-motivated set of gold standard responses
d) assessment specifications (...)
e) classify tasks within a continuum between non-communicative learning and authentic communication.

**Figure:** Communicative vs. non-communicative task continuum
AASF: Automatic Assessment Specifications Framework

Specifications for Automatic Linguistic Analysis

- what to analyse,
- how to code it
- required NLP modules/tasks

Specifications for the Feedback Generation Logic

- use linguistic codes to make hypotheses on the correctness of learner responses
- provide “canned” feedback messages
Analysing ICALL tasks

- Task type I: Create a customer satisfaction questionnaire
- Task type II: Describe the organisation of company using a chart
- Task type III: Write an email to register for a course in your own company
- Task type IV: Escribe una carta a Chupa Chups dándoles tu opinión sobre Smint
Empirical validation of design-based specifications I

1. Do learner responses correlate with pedagogical goals?
   - Response length: shorter vs. open/longer response
   - Relationship between input and response
     - Focusing on form makes prediction easier, but variation occurs

2. What kinds of variation can we observe?
   - Well- and ill-formed variation occur frequently in all tasks
   - Well-formed variation is higher in more open tasks
Empirical validation of design-based specifications II

3. Are design specifications useful?
   - Specifications account for 50% of actual response fragments

4. Can learner responses help improve NLP strategies?
   - Frequent patterns emerge from observed variation

⇒ Argues for the combination of design-monitored corpus driven approaches
Four task types in the middle ground

- Decontextualised grammar drills
- Type I: guided short-response
- Type II: picture description
- Type III: email in professional context
- Type IV: opinion letter

COMMUNICATIVE // COMPLEX

MIDDLE GROUND

 Essays with individualised topics
Four task types in the middle ground

- Decontextualised grammar drills
- Type I: guided short-response
- Short RC
- Type II: picture description
- Type III: email in professional context
- Type IV: opinion letter
- Essays with individualised topics

Non-communicative Learning → Pre-communicative Learning → Communicative Language Practice → Structured Communication → Authentic Communication
Introduction

Background

ICALL tasks – Where FLTL meets NLP

Enabling teachers to author ICALL activities
  - Customisation of an NLP-based feedback generation strategy
  - Integrating ICALL in secondary education environments

Conclusions
A context of application for an authoring tool
A customisable architecture for NLP-based assessment

Customisation of an NLP-based feedback generation strategy

LEARNER RESPONSE → WORD AND TEXT SEGMENTATION → MORPHOLOGICAL ANALYSER → MORPHOSYNTACTIC DISAMBIGUATION

ASSESSMENT WITH GENERAL NLP RESOURCES → CONTEXT SENSITIVE SPELL CHECKING → SPELL CHECKING

INFORMATION EXTRACTION MODULE(S) → FEEDBACK GENERATION MODULE(S) → ASSESSMENT WITH TASK SPECIFIC NLP RESOURCES

CUSTOMISE WITH RESPONSE SPECIFICATIONS
Customisation of an NLP-based feedback generation strategy

Response Specification Language

- a means for content developers to provide a set of correct responses
- contains the criteria for correctness, its thematic and linguistic contents
- can be expanded to model a larger set of responses
Customisation of an NLP-based feedback generation strategy

Linguistic- and surface-based response expansion

(a) Substitution I

(b) Substitution II

(c) Addition

(d) Reordering
ReSS: Response Specification Scheme

1. Identify things to be said ⇒ Response Components
ReSS: Response Specification Scheme

1. Identify things to be said ⇒ Response Components
2. Classify sentence fragments into Response Components ⇒ Variants
Customisation of an NLP-based feedback generation strategy

ReSS: Response Specification Scheme

1. Identify things to be said ⇒ Response Components
2. Classify sentence fragments into Response Components ⇒ Variants
3. Identify combinatorial restrictions between Variants
ReSS: Response Specification Scheme

1. Identify things to be said ⇒ Response Components
2. Classify sentence fragments into Response Components ⇒ Variants
3. Identify combinatorial restrictions between Variants
4. Identify optional/alternative elements in response fragments (different ways of expressing the same thing)
ReSS: Response Specification Scheme

1. Identify things to be said \(\Rightarrow\) Response Components
2. Classify sentence fragments into Response Components \(\Rightarrow\) Variants
3. Identify combinatorial restrictions between Variants
4. Identify optional/alternative elements in response fragments (different ways of expressing the same thing)
5. State orders in which variants can be combined \(\Rightarrow\) RC Sequences
ReSS applied: How did E.T. learn English?

**Response Component A**
- E.T./it learnt to speak English/it
- it/E.T. learnt it/English

**Response Component B**
- by repeating the words
- by repeating what

**Response Component C**
- E.T./it heard
- E.T./it listened to
- that E.T./it heard
- that E.T./it listened to

Gertie/Elliot’s sister repeated
- Gertie/Elliot’s sister said

while Elliot’s sister watched Sesame Street

**RCS7** < A1?, B2, C1, D1 >

**RCS8** < A1?, B2, C1, D2 >

**RCS9** < A1?, B2, C3, D2 >

**RCS10** < A1?, B2, C3, D3 >
Instructional conditions

- Blended learning approach
- Use of individual learning activities to support a syllabus
- Generation of formative feedback

Technical conditions

- Be compatible with Moodle
Integration of an authoring tool in secondary education environments

**Diagram:**
- **ATTACK (GUI)**
  - Response Components
  - Response Component Sequences
  - Pre-defined Error Patterns
- **ATTACK (Back-end)**
  - Process
  - Content Analyser
  - Specific Lexicon
- **ATAP (Back-end)**
  - Process
  - Global Response Checker
  - Customised Error Checker
  - Feedback Generator
- **Response Specification Scheme**
- **Customisation of Response Evaluation Strategy**
- **Resources for Learner Response Evaluation**
Experiment characteristics

Participants

- Teacher 1: *Science* in 3rd year
- Teachers 2 and 3: EFL in 1st/2nd year

Procedure

- Teacher training
- Material creation
- Use of materials

Resulting materials

- T1: a lesson plan including lectures, lab experiments and PC-lab activities
- T2/T3: 5 activities to be used as supplementary material
How good was the feedback?

Analysing results for three different groups, one for each teacher.

### Table: Feedback at the level of spelling and grammar.

<table>
<thead>
<tr>
<th></th>
<th>False</th>
<th>True</th>
<th>Bad</th>
<th>Conn. F.</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2</td>
<td>97</td>
<td>1</td>
<td>0</td>
<td>1221</td>
</tr>
</tbody>
</table>

### Table: Feedback at the level of task-specific language and content.

<table>
<thead>
<tr>
<th></th>
<th>False</th>
<th>True</th>
<th>Bad</th>
<th>Not sub.</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1ESO-A1T2/3</td>
<td>0</td>
<td>76</td>
<td>0</td>
<td>24</td>
<td>233</td>
</tr>
<tr>
<td>2ESO-A1T2/3</td>
<td>0</td>
<td>63</td>
<td>0</td>
<td>37</td>
<td>318</td>
</tr>
<tr>
<td>3ESO-A1T1</td>
<td>0</td>
<td>60</td>
<td>13</td>
<td>27</td>
<td>993</td>
</tr>
</tbody>
</table>

Table: Feedback at the level of task-specific language and content.
Did learners benefit from feedback?

![Bar chart showing learner feedback across different tasks.](image)

**Spelling and grammar**

Learner perception: 53% found it very useful, 38% useful, and 8% not very useful.

**Task-specific**

Learner perception: 53% found it very useful, 38% useful, and 8% not very useful.
The teacher’s perspective

1. System’s feedback: positive impact on learner’s competences

2. Effects on teaching process:
   - easier integration of content and language
   - insight gain on the learner’s task completion procedures

3. Added values: positive influence, learner engagement, and learner challenging the computer (game), and greater learner autonomy.

4. Material creation process:
   - response specification
   - difficulty in determining NLP capabilities
How often did the teacher and the system agree?

**2B-A1T2/3 - Grammar and spell checking**

<table>
<thead>
<tr>
<th>Q1-3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
</tr>
</thead>
<tbody>
<tr>
<td>138</td>
<td>47</td>
<td>44</td>
<td>23</td>
<td>19</td>
<td>24</td>
</tr>
</tbody>
</table>

- **Teacher disagrees**: Green
- **Teacher agrees**: Yellow

**2B-A1T2/3 - Language and content checking**

<table>
<thead>
<tr>
<th>Q1-3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td>26</td>
<td>26</td>
<td>11</td>
<td>9</td>
<td>13</td>
</tr>
</tbody>
</table>

- **Teacher disagrees**: Green
- **Teacher agrees**: Yellow
Discussion I

Material creation process:

- Integration of materials was coherent with teacher/learner expectations
- Input data to support learners was used variably but did not seem to dramatically impact on system performance
- The ReSS proved to be a natural way of specifying responses
- Teachers ended up with a considerable understanding of what an NLP-enhanced authoring tool allowed them to do
Use of materials in class:

- T1 made use of the response update functionalities to improve response specification between classes
- Generally learners did take advantage of the system’s feedback
Discussion III

System limitations: Variation in learner responses

(1) I start LESSONS at nine o’clock.
   I start school at nine o’clock.

(2) I brush my hair and my teeth at a quarter to nine.
   I brush my hair at a quarter to nine., or
   I brush my teeth at a quarter to nine.
Automatic analysis of language vs. feedback generation

The enumeration contains part of the expected elements, but one of them is incorrect.
1. Introduction

2. Background

3. ICALL tasks – Where FLTL meets NLP

4. Enabling teachers to author ICALL activities

5. Conclusions
   - Contributions
   - Future work
Connecting TBLT and NLP principles

1. A methodology to guarantee a pedagogically and computationally principled design of ICALL tasks:
   - Task Analysis Framework (TAF)
   - Response Interpretation Framework (RIF)
   - Automatic Analysis Interpretation Framework (AASF)

2. Application of the methodology to TBI CALL materials

3. Analysis of learner responses providing linguistic evidence: informs on task’s complexity and NLP capabilities

4. Further characterisation of middle processing ground
NLP as an enabling technology for teachers

1. Methodology and technology for teachers to autonomously design, implement and use ICALL activities in class
2. Evaluation an ICALL authoring tool/methodology in secondary school instruction settings
3. Inclusion of teacher and learner perspective through the investigation
Future work

Thesis-related short term research

1. Collaborate with researchers in language teaching/acquisition to further assess the complexity of learning tasks (Robinson, 2011).

2. NLP as an enabling technology
   - Simplify/Enhance the response specification process
   - Corpus-driven response expansion process
   - Further customisation of assessment functionalities

3. Contrastive analysis of teacher assessment vs. computer assessment
More comprehensive methodology for the development and analysis of CALL materials

Design and development of materials taking into account pedagogy and technology from the beginning

Design and evaluate practices to help teachers/learners achieve a greater autonomy

(Colpaert, 2006; Schulze, 2008)
Thanks for your attention!
Bibliography I


Bibliography II


