What results give which strategies in error annotation

Ana Díaz-Negrillo & Salvador Valera
University of Jaén
Spain

HUM2007-30559-E/FILO

AALL’09
March 10th-11th 2009
Objectives

- Share with you a number of issues regarding identification and classification of errors in learner corpora
- Encourage discussion
- Favour standardisation
Materials

• Examples from
  – SPICLE: Spanish component in ICLE (200,376 words)
    • German component in ICLE
  – NOCE: English learner corpus by Spanish speakers (269,237 words)

• Error taggers covered
  – Louvain tagger: ICLE (Dagneaux et al. 1998)
  – NICT JLE tagger, English by Japanese learners (Izumi et al. 2005)
  – EARS, English by Spanish learners (Díaz Negrillo 2007)
  – Kölmyr (2003) To Err is Human, English by Swedish learners
  – FRIDA, learner French (Granger 2003)
  – FALKO, learner German (Lüdeling et al. 2005)
Learner corpus researchers’ claim: the weaknesses of previous data collections can now be overcome thanks to the general features of learner corpora (Nesselhauf 2005: 41):

- large size, representativeness, machine-readable format, contextualization of occurrences

- Error tagging are largely determined by the nature of learner corpora

- Agreement on a most adequate annotation policy is very much in need
Background

Error tagging in learner corpora:

What for?

- **Mark** what otherwise would be inaccessible
- **Group** errors of the same nature
- **Provide** descriptions of errors (CALL, FLT, SLA)

What type of descriptions?

- **Objective, systematic and useful**
Contents

- Error identification
- Error classification
- Error description
Error identification

Issue 1. What is an error?

Coder (1973: 260): “those features of the learner’s utterances which differ from those of any native speaker”

Lennon (1991: 182): “a linguistic form or combination of forms which, in the same context, and under similar conditions of production would, in all likelihood, not be produced by the native speaker counterparts”

Error = various types of non-native performance (grammatical, semantic, syntactic deviances and whatever sounds unnatural to the NS)
1) One of these reasons is [...] GR-1-A-EN-003-X

2) I decided to study “Filología inglesa” in Granada. The first problem was the inscription [...] GR-1-A-EN-021-F

3) [...] when you travel somewhere, if you don’t speak the first language of there, you can always [...] GR-1-A-EN-0-40-X

4) An explicit example to take in the hand is how one kills a snake. ICLE-GE-SAL-0014.4
Error identification

• Error taggers mostly cover grammatical and lexical errors while description of other instances of unnatural language do not attract as much attention.

• Grammatical but unnatural learner language is still non-native and might be equally described.
  — It’s most common among advanced learners.
Error identification

Corder (1973): mistakes vs. errors
- mistakes (NS and NNS), can be self-corrected
- errors (NS), can’t be self-corrected, interest of SLA research

Issue 2. Error = lack of knowledge?

5) [...] these eople who wanted to improve [...] ICLE-SP-UCM-0019.3

6) [...] something than in my opinion could be very interesting. [...] GR-1-A-EN-095-F
Error identification

How do people deal with this?

- ‘Typo’: FRIDA
- No distinction mistake vs. error because it is difficult to decide: Kölmyr (2003)
- A distinction is made at the explanation level: FALKO (Lüdeling et al. 2005)

Mistakes vs. errors. You can:
- Ask writers
- Check whether the occurrence is systematic / recursive
- Inter-rater reliability tests
- ....
Error identification

Issue 2. Mistake vs. error?

9) [...] something than in my opinion could be very interesting.

GR-1-A-EN-095-F

Matching performance to competence in learner corpora might not be so straightforward
Error identification

Issue 3. How many errors are there?

7) Mary Wollstonecraft *write* in 1792 *Vindications of the Rights of Women*. <ICLE-SP-UCM-0040.4>

Tense? 3rd person singular? Both?

8) [...] the oil, butter, milk that they *apperence* like very healthy products, [...] GR-1-B-EN-044-Z

Word class? Spelling? Both?

---

Kölmyr (2003): Minimum Correction Principle (MCP)

EARS: underlying errors are tagged.

Describing learners’ performance or correcting their writing?
Error description

Linguistic description of errors

- Issue 4: Basis of description
- Issue 5: Borderline cases
Error description

Issue 4. Basis of description:
targeted form? learner’s performance?

9) The main feature of a campus like Louvain-la-Neuve is (GA)
the $its$ conviviality
10) Nowadays the mobile phone is using by children. GR-
    1-B-EN-062-Y

Voice, passive? active?
Aspect, continuous? non-continuous?

FRIDA, Louvain, EARS:
- there might be various targeted utterances, so describing
  objective observable data might encourage objective and
  consistent annotation process
Error description

**Issue 4. Basis of description = ?**

10) our parents, teachers, etc. who want that we get our *mets* and we don’t want disappoint them. GR-1-A-EN-066-F

11) mobiles call the attention of *people* [...] GR-1-B-EN-062-Y

12) [...] they forgot the primary use of *them* [...] GR-1-B-EN-027-Y

If what is observed is described:

- How do describe cases that do not correspond to English forms? (ex. 10, verb)

- This might result in that various targeted possibilities might fall under the same tag/description (ex. 11 and 12)
Issue 5. Borderline cases

13) Then your parents don’t **say** you: you must eat more [...] GR-1-A-EN-029-Y

Omission of the preposition? Lexical misselection? Both?

14) They could affect the **hear** basically [...] GR-1-B-EN-061-Y

Word class? Spelling?

**EARS:**
Describing objective observable data, i.e. learners’ choices, seems to encourage objective and consistent annotations
Error classification

- Issue 6. Error information types
- Issue 7. Delicacy of description
Error classification

Issue 6. Error information types

Louvain, FRIDA, NICT JLE, EARS
Level: grammar, spelling, lexis, etc.
Unit: POS, diacritics, comma, etc.
System: tense, collocation, derivation, etc.

FALKO:
Error explanation

Target modification taxonomy (omission, substitution, ordering, intrusion)?
Execution errors (*childs) vs. errors in category selection (*one children)?

Description of errors according to more than one information type enables a variety of searches and purposes
However, it is a slow and tedious process
Error classification

Issue 7. Detail of description

Most taggers are generic (Louvain, NICT JLE)
A fine-grained error tagger: EARS

15) I decided to study “Filología inglesa” in Granada. The first problem was the inscription [...] GR-1-A-EN-021-F
16) [...] it’s an objective that I have to get. GR-1-A-EN-002-F
17) [...] took my things...phone, money, bonobus and [...] GR-1-A-EN-013-F
18) I tried to entry, [...] GR-1-A-EN-013-F
19) [...] is useless [...]. <ICLE-SP-UCM-0019.3>

20) Nowadays the mobile phone is using by children. GR-1-B-EN-062-Y

Voice, passive Aspect, continuous
Issue 7. Detail of description

Most taggers are generic (Louvain, NICT JLE)
  error tagging is faster
  seem to enable reusability
  may require further categorization from the end user
Conclusions

1. Because of the specific nature of learner corpora, error tagging calls for its own policy, which may differ from traditional SLA research
   - subjects are not available to the researcher
   - observable data is all that there is left
2. Focusing on observable data might enforce consistent, objective annotations
3. Annotation policy and, consequently, results are determined by the purposes of error tagging
   - correction or description of learner performance?
4. While most error taggers focus on grammatical errors there is still an area which remains largely unexplored: unnatural language
   - categories will rise when cases of this nature are grouped together and studied carefully
References


What results give which strategies in error annotation

Ana Díaz-Negrillo & Salvador Valera
University of Jaén
Spain

HUM2007-30559-E/FILO

AALL’09
March 10th-11th 2009