

ERG adaptation for grammar-checking with ESL learners

Dan Flickinger

CSLI, Stanford University

DELPH-IN Summit 2019

Cambridge

15 July 2019

Background

- Automated error detection in student writing
- Error-specific recommendations for repair
- High precision and high recall



Current applications

- Language Arts & Writing course for McGraw-Hill Education
In U.S. classrooms for past ten years
10,000 students, primary school, mostly native English
15 million sentences composed
1 million paragraphs (narrative, descriptive, opinion)
100K three-paragraph essays



Current applications

- Language Arts & Writing course for McGraw-Hill Education
In U.S. classrooms for past ten years
10,000 students, primary school, mostly native English
15 million sentences composed
1 million paragraphs (narrative, descriptive, opinion)
100K three-paragraph essays
- On-line test preparation for ESL exam, with Beijing publisher
High school ESL students
Single-sentence translation: 100K high school students
Short essay composition: 1.5 million middle and high school



Single-sentence translation

- Present a sentence in Mandarin for translation
- Evaluate for spelling, punctuation, capitalization, grammar
- If mechanically well-formed, evaluate semantics
- Invite the student to correct errors and re-submit



An example

Reference translation:

My family and I plan to travel in a European country during the winter holiday to see the incredible scenery there.

Student answer:

My family and I are going to travel to an Europe country to see the beautiful views in winter holiday.



Data sets for this task

- 30 Chinese sentences for translation (one per day for a month)
- 30 reference translations authored for each Chinese sentence
- 50 teacher-authored simulated student responses for each one
- 100 student-authored responses for each one
- Manual evaluation of system response to each of these 5000+ English sentences



Infrastructure

- Existing app for improving pronunciation, reading English aloud
- Addition of this translation task as new feature
- Servers running ACE parser with customized ERG
- Mapping from grammar's 'error codes' to instructional responses
- Conversion from MRS to EDS, for graph-matching with references



Status of pilot

- First cycle of quality assurance
 - For each of 900 reference translations, good top-ranked parse
 - For 3800 unique elicited translations, 2300 judged correct
 - Of the 1500 wrongly analyzed, majority are due to semantics
- Second cycle of QA now in process, after Stanford improvements
 - Aiming for precision above 95%.



ERG customizations

- Additions

Mal-rules such as for missing article, or adverb as prenom adj

Mal-entries such as for wrong complement type (e.g. NP vs PP)

- Revisions

Rule for NP-N compounds modified to block *Kim and me skating*

Lexical entry for noun “China” blocked in *China people*

- Blocking

Construction for subjunctive inversion *Were we to win, ...*

Lexical entry for verb “cream”



Summary of ERG customizations

- 250 error types, triggered by 50 mal-rules and 1300 mal-entries
- 60 blocked constructions
- 70 modified constructions
- 600 blocked lexical entries
- 400 modified lexical entries

