# Analyzing an English learner corpus and system responses

# **Dan Flickinger**

CSLI, Stanford University

**DELPH-IN Summit 2020** 

16 July 2020

# Data collection setting: Up366 in Beijing

- Commercial product for grammar-checking in use since Feb 2020
- 2,000 middle/high school students in China learning English
- Each day, one Chinese sentence to translate into English
- ACE with ERG + mal-rules running on servers
  See my summit talks from 2013 and 2019 for details
- Evaluate spelling, punctuation, capitalization and grammar



## **Sentence-a-Day examples**

Presented to the student:

在周日,我经常和我妈妈一起去我家附近的一家购物中心购物。

Sample correct translation:

I usually go shopping with my mother in a shopping mall near my home on Sunday.

Examples of incorrect student answers:

- On *sunday*, *l* often go shopping in a shopping mall near my home with my mother.
- On *Sunday,I* often go shopping in a shopping center around our house with my mother.
- On Sunday, I often go shopping *with* in a shopping center near my home with my mother.
- On Sunday, I often go shopping *in shopping center* around my house with my mother.
- I often go shopping with my mom on the near



#### **Evaluating system responses**

Four possible coarse-grained pairings:

- Correct answer, rightly evaluated as correct
- Correct answer, wrongly evaluated as incorrect
- Incorrect answer, wrongly evaluated as correct
- Incorrect answer, rightly evaluated as incorrect



#### **Evaluating system responses**

Four possible coarse-grained pairings:

- CC Correct answer, rightly evaluated as correct
- CX Correct answer, wrongly evaluated as incorrect
- XC Incorrect answer, wrongly evaluated as correct
- XX Incorrect answer, rightly evaluated as incorrect

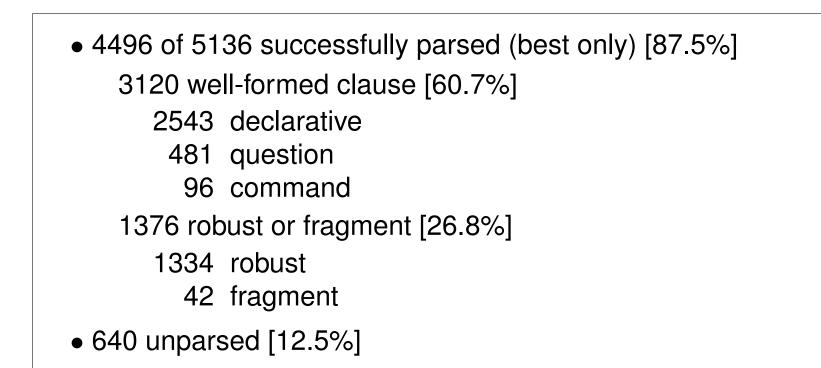


## **Student data from Month 3**

- 51,360 student answers over the 27 days
- Sampled every tenth sentence from time-ordered spreadsheet
- Began manual annotation of this 5136-sentence sample Average sentence length: 12.6 tokens Unique answers: 3635 of the 5136



# Parsing with ACE + ERG-mal





## Manually evaluating system responses

- CC ? of 3120 (correct, reported correct)
- XC ? of 3120 (error, reported correct)
- **CX** 50 of 5136 (correct, reported with error(s)) [1%] 10 of 640 unparsed

Many people pursue happiness [ all their lives. ]

1 of 42 fragments

Establishing [ the natural reserve is a good way to provide more living space for these wild animals. ]

39 of 1334 robust sentence analyses

XX 1966 of 5136 (error, reported as error) [38.3%]
 1295 of 1334 robust sentence analyses
 630 of 640 unparsed

41 of 42 fragments



### **Interim report**

- False reports of errors in less than one percent of answers
- Correct reports of errors for 38.3% of answers
- Don't yet know how often the system failed to detect errors
- Don't yet know how effective the error feedback is for learning

