Automated Writing Assistance: Grammar Checking and Beyond
Topic 4: Handling ESL Errors

Robert Dale
Centre for Language Technology
Macquarie University

From the Pasadena Star-News 26th Aug 2011

Robot teaches English as Second Language

Janette Williams, Staff Writer
Posted: 08/26/2011 06:33:10 PM PDT

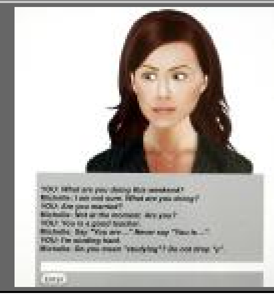
PASADENA - Say "How do you do" to Mike and Michelle, face-to-face tutors for English learners.

They'll correct your grammar, answer questions, converse on a variety of topics, be there 24/7, and won't charge a dime.

And they're doing very well, thank-you.

The on-screen "English Tutor" interactive robots and their creator, adjunct Professor Ron Chang Lee of Pasadena City College, are heading to England's Exeter University in October as one of four finalists in the 2011 Loebner Prize for Artificial Intelligence.

Click photo to enlarge



Pasadena City College adjunct assistant Professor Ron Chang Lee displays Michelle, his English as...

1 2 »

"I always wanted to create something to help students, like a tutor," said Lee, who has taught English as a Second Language at PCC since 1991.

"It's a talking robot, so (students) are not afraid of asking anything," Lee said. Interacting with a robot is less intimidating for ESL students than conversing with a professor, he said.

"A community college like PCC has many international students, and first they have to listen to English, to communicate with their professors," he said. "So the first year they have to take ESL."

Just 15 minutes a day with Mike or Michelle can really help, he said.

Outline

- Background
- Article Errors
- Preposition Errors
- Other ESL Problems
- Conclusions

Terminology

- ESL = English as a Second Language
  - Refers to non-native speakers living and speaking in a predominantly English-speaking environment
- EFL = English as a Foreign Language
  - Refers to non-native speakers studying and learning English in a non-English speaking country
- We'll generally use the term ESL to refer to both
- Apologies that this is mostly about <u>E</u>SL – there's less work in other languages …

The Problem

- Lots of people want to speak English: it is the most commonly studied second language
- Over 1 billion people speak English as a second or a foreign language
- Existing grammar checking tools are not, so far, tailored to the needs of ESL learners

ESL Errors Are Different: Bolt [1992]

- Bolt tested seven grammar-checking programs of the time against 35 sentences containing ESL errors
- Looked at from the perspective of a learner of English at a fairly low level of competence
- Conclusions:
  - ‘all of these programs fail in terms of the criteria that have been used.’
  - Expectations are encouraged that cannot be fulfilled
  - Silence on the part of a program suggests everything is ok

ESL Errors Are Different: Donahue [2001] vs Connors + Lundsford [1988]

Error	US	ESL
No comma after introductory element	1	negligible
Vague pronoun reference	2	negligible
No comma in compound sentence	3	12
Wrong word	4	2
No comma in nonrestrictive element	5	negligible
Wrong or missing inflected ends	6	6
Wrong or missing preposition	7	5
Comma splice	8	1
Possessive apostrophe error	9	negligible
Tense shift	10	negligible
Unnecessary shift in person	11	15
Sentence fragment	12	7
Wrong tense or verb form	13	4
Subject-verb agreement	14	11
Lack of comma in a series	15	negligible
Pronoun agreement error	16	negligible
Unnecessary commas with restrictive relative pronouns	17	negligible
Run on, fused sentences	18	8
Dangling, misplaced modifier	19	negligible
Its, it's confusion	20	negligible

ESL Errors Are Different

Error	US	ESL
Missing words	negligible	3
Capitalization	negligible	9
Wrong pronoun	negligible	16
a, an confusion	negligible	14
Missing article	negligible	17
Wrong verb form	negligible	10
No comma before etc.	negligible	13

- **Half of the ten most frequent error types made by native speakers are negligible in the writing of the ESL population**

Errors in the Cambridge Learners Corpus

Rank	Error Type	Prop	Example sentence
1	Content word choice error	0.199	We need to deliver the merchandise on a daily **base/basis*.
2	Preposition error	0.134	Our society is developing **in/at* high speed.
3	Determiner error	0.117	We must try our best to avoid **the/a* shortage of fresh water.
4	Comma error	0.093	However, */, I'll meet you later.
5	Inflectional morphology	0.074	The women **weared/wore* long dresses.
6	Wrong verb tense	0.067	I look forward to **see/seeing* you.
7	Derivational morphology	0.049	It has already been **arrangement/arranged*.
8	Pronoun	0.042	I want to make **me/myself* fit.
9	Agreement error	0.040	I **were/was* in my house.
10	Run-on Sentence	0.040	The deliver documents to them *they* provide fast service.
11	Idiomatic Collocation and word order	0.039	The latest issue **the magazine of/of the magazine* …
12	Confused words	0.019	I want to see the **personal/personnel* manager.
13	Conjunction error	0.017	I want to see you **and/so* that you can help me.
14	Words split with a space or joined	0.014	I organize sports **everyday/every day*. It is also my **life style/lifestyle*.
15	Apostrophe error (not including *it/it's* confusions)	0.013	We are all **sport's/sports* lovers.
16	Hyphenation error	0.013	It is a nourishing **low cost/low-cost* meal.
17	Sentence fragment or two sentences that are joined	0.008	I'm going to get another one *. *Because/because* the old one broke.
18	Quantifier error	0.007	It doesn't give them too **much/many* problems.
19	Other punctuation error	0.004	When are you leaving *./?
20	Negation formation	0.001	We **have not/do not have* any time.

Common ESL Errors

- The most difficult aspects of English for ESL learners are:
  – Definite and indefinite articles
  – Prepositions
- Together these account for 20–50% of grammar and usage errors
- [The elephant in the room: spelling errors are much more common, and incorrect word choice is as problematic as article and preposition errors.]

Article Errors in the CLC by L1

L1	Has Articles	Proportion
Russian	No	0.186
Korean	No	0.176
Japanese	No	0.159
Chinese	No	0.125
Greek	Yes	0.087
French	Yes	0.081
Spanish	Yes	0.070
German	Yes	0.053

Proportion of sentences with one or more article errors

Preposition Errors in the CLC by L1

L1	Proportion
Greek	0.149
Spanish	0.139
Korean	0.128
Chinese	0.122
French	0.121
Japanese	0.118
German	0.100
Russian	0.095

Proportion of sentences with one or more preposition errors

The Impact of L1 on ESL Errors

- Learning will be difficult if the L1 has no close equivalent for a feature:
  - Native speakers of Japanese and Russian will have particular difficulty mastering the use of articles.
- Learning will be facilitated if the L1 has an equivalent feature:
  - Native speakers of French or German should find the English article system relatively easy to learn.

A Note on Data

- The field has been hamstrung by the privately held nature of many learner corpora
- Two welcome changes:
  - The NUS Corpus of Learner English
  - The Cambridge Learner Corpus FCE Dataset
- Also the much smaller HOO dataset

NUCLE: The NUS Corpus of Learner English

- 1400 essays written by University students at the National University of Singapore
- Over 1M words annotated with error tags and corrections
- See http://nlp.comp.nus.edu.sg/corpora

NUCLE: The NUS Corpus of Learner English

Standoff annotation:

```
<MISTAKE start_par="4" start_off="194" end_par="4" end_off="195">
   <TYPE>ArtOrDet</TYPE>
   <CORRECTION>an</CORRECTION>
</MISTAKE>
```

The CLC FCE Dataset

- A set of 1,244 exam scripts written by candidates sitting the Cambridge ESOL First Certificate in English (FCE) examination in 2000 and 2001
- Annotated with errors and corrections
- A subset of the much larger 30M-word Cambridge Learner Corpus
- See http://ilexir.co.uk/applications/clc-fce-dataset/

The CLC FCE Dataset

Inline annotation:

- Because <NS type="UQ"><i>all</i></NS> students in <NS type="MD"><c>the</c></NS> English class are from all over the world …

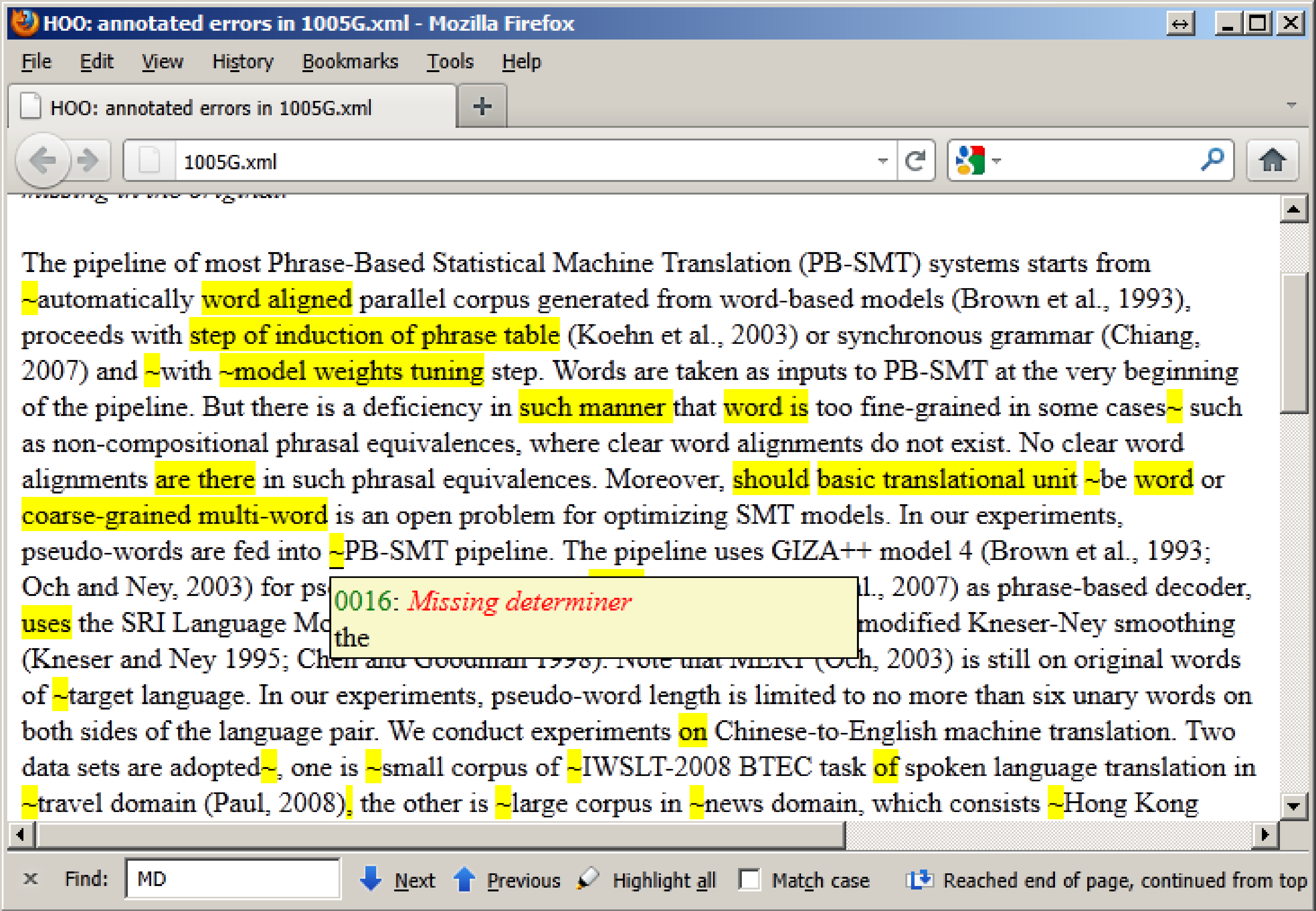
The HOO Dataset

- HOO – Helping Our Own – aims to marshall NLP technology to help non-native speakers write ACL papers
- Very small corpus (~36K words) annotated with errors and corrections
- Evaluation software also freely available
- See http://www.clt.mq.edu.au/research/projects/hoo/

The HOO Dataset

Stand-off and inline annotation both available:

- In our experiments, pseudo-words are fed into <edit type="MD"><empty/><corrections><correction>the</correction></corrections></edit> PB-SMT pipeline.
- <edit index="1005-0016" type="MD" start="871" end="871" >
  <original><empty/></original>
  <corrections>
  <correction>the </correction>
  </corrections>
  </edit>


HOO: annotated errors in 1005G.xml - Mozilla Firefox
File Edit View History Bookmarks Tools Help
HOO: annotated errors in 1005G.xml
1005G.xml
The pipeline of most Phrase-Based Statistical Machine Translation (PB-SMT) systems starts from ~automatically word aligned parallel corpus generated from word-based models (Brown et al., 1993), proceeds with step of induction of phrase table (Koehn et al., 2003) or synchronous grammar (Chiang, 2007) and ~with ~model weights tuning step. Words are taken as inputs to PB-SMT at the very beginning of the pipeline. But there is a deficiency in such manner that word is too fine-grained in some cases~ such as non-compositional phrasal equivalences, where clear word alignments do not exist. No clear word alignments are there in such phrasal equivalences. Moreover, should basic translational unit ~be word or coarse-grained multi-word is an open problem for optimizing SMT models. In our experiments, pseudo-words are fed into ~PB-SMT pipeline. The pipeline uses GIZA++ model 4 (Brown et al., 1993;
Och and Ney, 2003) for ps
l., 2007) as phrase-based decoder,
uses the SRI Language Mo
modified Kneser-Ney smoothing
(Kneser and Ney 1995; Ch
ch, 2003) is still on original words
0016: Missing determiner
the
of ~target language. In our experiments, pseudo-word length is limited to no more than six unary words on both sides of the language pair. We conduct experiments on Chinese-to-English machine translation. Two data sets are adopted~, one is ~small corpus of ~IWSLT-2008 BTEC task of spoken language translation in ~travel domain (Paul, 2008), the other is ~large corpus in ~news domain, which consists ~Hong Kong
Find: MD
Next
Previous
Highlight all
Match case
Reached end of page, continued from top

Outline

- Background
- Article Errors
- Preposition Errors
- Other ESL Problems
- Conclusions

Article Errors

- The Problem
- Early Rule-based Approaches
- Knight and Chandler [1994]
- Han et al [2006]
- De Felice and Pulman [2008]

Why is Article Choice Hard?

- Basic problem for speakers of languages that do not use articles:
  - choose between a/an, the, and the null determiner
- The bottom line: it comes down to context
  - I was eating a cake.
  - I was eating the cake.
  - I was eating cake.

Features Impacting Article Choice: Countability

- Count nouns take determiners:
  - I read the paper yesterday.
- Mass nouns don't take determiners:
  - We generally write on paper.
- But the universal grinder and the universal packager [Pelletier 1975] are always available:
  - There was dog all over the road.
  - Could we have just one rice please?

Features Impacting Article Choice: Countability

- Semi-idiomatic forms:
  - I looked him in the eye.
  - *I looked him in an eye.

Features Impacting Article Choice: Syntactic Context

- ✓ I have knowledge.
- × I have a knowledge.
- ✓ I have knowledge of this.
- × I have a knowledge of this.
- ✓ I have a knowledge of English.

Features Impacting Article Choice: Discourse Factors

- Stereotypically, entities are introduced into a discourse using an indefinite determiner and subsequently referred to using a definite determiner
  - I saw <u>a man</u> at the bus stop. … <u>The man</u> was crying.
- But not always:
  - A bus turned the corner. <u>The driver</u> was crying.
  - I went to <u>the beach</u> yesterday.

Features Impacting Article Choice: World Knowledge

- He bought a Honda.
- He bought Honda.

Article Errors

- The Problem
- Early Rule-based Approaches
- Knight and Chandler [1994]
- Han et al [2006]
- De Felice and Pulman [2008]

Early Work: Article Insertion in Machine Translation

- The Problem:
  - Machine translation of languages like Japanese or Russian into English is difficult because the source language doesn't contain articles

Murata and Nagao [1993]: Hand-Crafted Rules

- When a noun is modified by a referential pronoun (KONO(this), SONO(its), …) then {indefinite(0, 0), definite(1, 2), generic(0, 0)}
- When a noun is accompanied by a particle (WA), and the predicate has past tense, then {indefinite(1, 0), definite(1, 3), generic(1, 1)}
- When a noun is accompanied by a particle (WA), and the predicate has present tense, then {indefinite(1, 0), definite(1, 2), generic(1, 3)}
- When a noun is accompanied by a particle HE(to), MADE(up to) or KARA(from), then {indefinite(1, 0), definite(1, 2), generic(1, 0)}
- …

Article Errors

- The Problem
- Early Rule-based Approaches
- Knight and Chandler [1994]
- Han et al [2006]
- De Felice and Pulman [2008]

Knight and Chandler [1994]:
A Data-Driven Method for Post-Editing

- General aim:
  - To build a post-editing tool that can fix errors made in a machine translation system
- Specific task:
  - Article insertion: *a*, *an* or *the*

Knight and Chandler [1994]: Before and After

Stelco Inc. said it plans to shut down three Toronto-area plants, moving their fastener operations to leased facility in Brantford, Ontario.

Company said fastener business “has been under severe cost pressure for some time.” Fasteners, nuts and bolts are sold to North American auto market.

Company spokesman declined to estimate impact of closures on earnings. He said new facility will employ 500 of existing 600 employees. Steelmaker employs about 16,000 people.

Stelco Inc. said it plans to shut down three Toronto-area plants, moving their fastener operations to a leased facility in Brantford, Ontario.

The company said the fastener business “has been under severe cost pressure for some time.” The fasteners, nuts and bolts are sold to the North American auto market.

A company spokesman declined to estimate the impact of the closures on earnings. He said the new facility will employ 500 of the existing 600 employees. The steelmaker employs about 16,000 people.

Knight and Chandler [1994]: The General Idea

The steps:

- Take newspaper-quality English text
- Remove articles
- Re-insert automatically
- Compare results with the original text

Assumptions:

- NPs are marked as singular or plural
- Locations of articles already marked so it's a binary choice between *the* and *a/an*.

Knight and Chandler [1994]: Baseline

- In 40Mb of Wall Street Journal text:

  a = 28.2%

  an = 4.6%

  the = 67.2%
- So 67% is a good lower-bound
- Upper-bound:
  - Human subjects performed with accuracy of 94%-96%

Knight and Chandler [1994]: Baselines

	Human	Machine
Random	50%	50%
Always guess *the*	67%	67%
Given core context NP	79-80%	
Given NP + 4 words	83-88%	?
Given full context	94-96%	

Knight and Chandler [1994]: Approach

- Characterize NPs via sets of features then use a build decision tree to classify
- Lexical features:
  - ‘word before the article is *triple*’
- Abstract features:
  - ‘word after the head noun is a past tense verb’
- 400k training examples and 30k features; features with less than 4 instances discarded

Knight and Chandler [1994]: Performance

- On 1600 trees for the 1600 most frequent head nouns (covering 77% of test instances):
  - 81% accuracy
- Guess *the* for the remaining 23% of test instances
  - 78% accuracy overall

Article Errors

- The Problem
- Early Rule-based Approaches
- Knight and Chandler [1994]
- Han et al [2006]
- De Felice and Pulman [2008]

Han et al [2006]: A MaxEnt Approach to Article Selection

- Basic Approach:
  - A maximum entropy classifier for selecting amongst *a/an*, *the* or the null determiner
  - Uses local context features such as words and PoS tags

Han et al [2006]: Contrasts with Earlier Work

- More varied training corpus: a range of genres
  - 721 text files, 31.5M words
  - 10^{th} thru 12^{th} grade reading level
- Much larger training corpus: 6 million NPs (15x larger)
  - Automatically PoS tagged + NP-chunked
- The use of a maximum entropy classifier

Han et al [2006]: Training Results

- 6M NPs in training set
- 390k features
- Baseline = 71.84% (frequency of null determiner)
- Four-fold cross validation
  - performance range 87.59% to 88.29%
  - Average 87.99%

Han et al [2006]: Effectiveness of Individual Features

Feature	% Correct
Word/PoS of all words in NP	80.41
Word/PoS of w(NP-1) + Head/PoS	77.98
Head/PoS	77.30
PoS of all words in NP	73.96
Word/PoS of w(NP+1)	72.97
Word/PoS of w(NP[1])	72.53
PoS of w(NP[1])	72.52
Word/PoS of w(NP-1)	72.30
PoS of Head	71.98
Head's Countability	71.85
Word/PoS of w(NP-2)	71.85
Default to null determiner	71.84

Han et al [2006]: Effectiveness of Individual Features

- Best feature: Word/PoS of all words in NP
  - Ok if you have a large enough corpus!
- Second best: W(NP-1) + Head
  - Eg ‘in summary’

Han et al [2006]:
Accuracy by Head Noun Type

Syntactic Type of Head	% Correct
Singular Noun	80.99
Plural Noun	85.02
Pronoun	99.66
Proper Noun, Singular	90.42
Proper Noun, Plural	82.05
Number	92.71
Demonstrative Pronoun	99.70
Other	97.81

Han et al [2006]: Accuracy as a Function of Training Set Size

#NPs in Training Set	% Correct
150000	83.49
300000	84.92
600000	85.75
1200000	86.59
2400000	87.27
4800000	87.92
6000000	87.99

Han et al [2006]: Applying the Model to TOEFL Essays

- Model retrained only on NPs with a common head noun
  - Baseline = frequency of null determiner = 54.40%
  - Training set kept at 6M instances by adding more data
  - Average accuracy = 83.00%
- Model applied to 668 TOEFL essays w 29759 NPs
  - Subset of NPs classified by two annotators
  - Agreement on 98% of cases with kappa = 0.86
  - One article error every 8 NPs

Han et al [2006]: Some Examples

Above all, I think it is good for students to share room with others.

- Human: missing *a* or *an*
- Classifier: 0.841 *a/an*; 0.143 *the*; 0.014 zero

Those excellent hitters began practicing the baseball when they were children, and dedicated a lot of time to become highly qualified.

- Human: superfluous determiner
- Classifier: 0.103 *a/an*; 0.016 *the*; 0.879 zero

Han et al [2006]: Results on TOEFL Essays

- 79% of errors in test set correctly detected
- Many false positives, so precision only 44%
- Decisions often borderline:
  - The books are assigned by <u>professors</u>.
  - Marked by annotators as correct, model predicts *the* (0.51) and null (0.49)

Han et al [2006]: Sources of Error

- Model performs poorly on decision between *a* and *the*
  - Probably due to the need for discourse information
- So, new feature: has the head noun appeared before, and if so, with what article?
  - No significant effect on performance
- Error analysis suggests this is due to more complex discourse behaviour:
  - A student will not learn if she hates the teacher.
  - … the possibilities that a scholarship would afford …

Article Errors

- The Problem
- Early Rule-based Approaches
- Knight and Chandler [1994]
- Han et al [2006]
- De Felice and Pulman [2008]

De Felice and Pulman [2008]: Richer Syntactic and Semantic Features

- Basic Approach:
  - As in Han et al [2006], a maximum entropy classifier for selecting amongst *a/an*, *the* or the null determiner
  - Use a richer set of syntactic and semantic features

De Felice and Pulman [2008]: Main Features

Feature	Value
Head Noun	‘apple’
Number	Singular
Noun Type	Count
Named Entity?	No
WordNet Category	Food, Plant
Prepositional Modification?	Yes, ‘on’
Object of Preposition?	No
Adjectival Modification?	Yes, ‘juicy’
Adjectival Grade	Superlative
POS±3	VV, DT, JJS, IN, DT, NN

Example: Pick <u>the juiciest apple on the tree</u>.

De Felice and Pulman [2008]: Additional Features

- Whether the noun is modified by a predeterminer, possessive, numeral and/or a relative clause
- Whether it is part of a ‘there is …’ phrase

De Felice and Pulman [2008]: Performance

- Trained on British National Corpus
  - 4,043,925 instances
- Test set of 305,264 BNC instances
- Baseline = 59.83% (choose null)
- Accuracy = 92.15%

De Felice and Pulman [2008]: Comparative Performance on L1 Data

Author	Accuracy
Baseline	59.83%
Han et al 2006	83.00%
Gamon et al 2008	86.07%
Turner and Charniak 2007	86.74%
De Felice and Pulman 2008	**92.15%**

De Felice and Pulman [2008]: Results on Individual Determiners

	% of Training Data	Precision	Recall
a	9.61% (388,476)	70.52%	53.50%
the	29.19% (1,180,435)	85.17%	91.51%
null	61.20% (2,475,014)	98.63%	98.79%

- **The indefinite determiner is less frequent and harder to learn**

De Felice and Pulman [2008]: Testing on L2 Text

- 3200 instances extracted from the CLC
  - 2000 correct
  - 1200 incorrect
- Accuracy on correct instances: 92.2%
- Accuracy on incorrect instances: < 10%
- Most frequent incorrect usage is a missing determiner
  - Model behaviour influenced by skew in training data
- Also problems in extracting NLP features from L2 data

Outline

- Background
- Article Errors
- Preposition Errors
- Other ESL Problems
- Conclusions

The Prevalence of Preposition Errors

L1	Proportion
Greek	0.149
Spanish	0.139
Korean	0.128
Chinese	0.122
French	0.121
Japanese	0.118
German	0.100
Russian	0.095

Proportion of sentences in the CLC with one or more preposition errors

Prepositions Have Many Roles in English

- They appear in adjuncts:
  - In total, I spent $64 million dollars.
- They mark the arguments of verbs:
  - I'll give ten cents to the next guy.
- They figure in phrasal verbs:
  - I ran away when I was ten.
- They play a part in idioms:
  - She talked down to him.

Negative Transfer

- Many prepositions have a most typical or frequent translation
  - Eg: *of* in English to *de* in French
- But for many prepositions there are multiple translational possibilities
  - ESL speakers can easily choose the wrong one
  - Eg: driving *in* a high speed

Prepositions in English

- English has over 100 prepositions, including some multiword prepositions and a small number of postpositions
- The 10 most frequent account for 82% of the errors in the CLC

Preposition Selection in Well-Formed Text

Citation	Approach	Training Corpus	Testing Corpus	Performance
Lee and Seneff (2006)	Parse Ranking, using Collins (1999)'s parser	10,369 Transcripts of flight reservation data	317 sentences from transcripts of flight reservation data	P=88%, R=78%
Chodorow et al. (2007)	Maximum Entropy Classifier and Heuristic Rules, token context, part-of-speech context, chunk information	SJM and MetaMetrics: 7M cases	MetaMetrics: 18.2K cases	69% accuracy
De Felice and Pulman (2007)	Voted Perceptron, part-of-speech context, parse information, semantic information	BNC: 10-fold xval	BNC subset, 10k sentences	76% accuracy
De Felice and Pulman (2008)	Maximum Entropy Classifier, part-of-speech context, parse information, semantic information	BNC: 9M cases	BNC: 536.2K cases	70% accuracy
Tetreault and Chodorow (2008b)	Maximum Entropy Classifier and Heuristic Rules, token context, part-of-speech context	SJM and MetaMetrics: 10M cases (plus NANTC and Encarta/Reuters)	WSJ, Encarta/Reuters (1.4M cases)	90% accuracy (WSJ), 79% accuracy (Encarta/Reuters)
Gamon et al. (2008)	Decision Tree and Language Model, token context, part-of-speech context	Encarta, Reuters, etc.	Encarta/Reuters 1.4M cases	combined accuracy =77% (presence/absence = 91%, choice=62%)
Bergsma et al. (2009)	Google N-gram corpus approach (log of counts)	NYT: 1M cases	NYT: 10K cases	75% accuracy

Preposition Error Detection on Learner Data

Citation	Approach	Training Corpus	Testing Corpus	Performance
Eeg-Olofsson and Knutt (2003)	Heuristic Rules, parse information	n/a	40 cases from Swedish learner essays	11/40 correct
Tetreault and Chodorow (2008a)	Maximum Entropy Classifier and Heuristic Rules, token context, part-of-speech context	SJM and MetaMetrics: 7M cases	TOEFL: 8.2K cases	P=84%, R=19%
De Felice and Pulman (2009)	Maximum Entropy Classifier, token context, part-of-speech context, semantic information	BNC: 9M cases	CLC: 1116 incorrect cases, 5753 correct cases	P=42% and R=35% on incorrect cases, accuracy 69% on correct cases
Hermet et al. (2008)	Web-counts method	WWW	133 French Learner sentences	70% accuracy on error correction task
Tetreault and Chodorow (2009)	Web-counts method (Region Counts Approach)	WWW	TOEFL: 518 cases for 5 constructions	n/a
Gamon (2010)	Maximum Entropy and LM, token context, part-of-speech context	2.5M sentences of well-formed text; LM (Gigaword); CLC for meta-classifier	CLC (208.7K sentences/19.7K errors)	Auto: P=35%, R=22%; Manual verification: 6K sentences, P=85%
Han et al. (2010)	Maximum Entropy, token context, part-of-speech context, parse information	Chungdahm: 978,000 error-annotated cases	Chungdahm: 1,000 cases	Detection: P=93%, R=15%; Correction: P=82%, R=13%

Outline

- Background
- Article Errors
- Preposition Errors
- Other ESL Problems
- Conclusions

Collocations

- Conventional combinations that are preferred over other equally syntactically and semantically valid combinations
  - Adj + Noun: *stiff breeze* vs *rigid breeze*
  - Verb + Noun: *hold an election* vs *make an election*
  - Noun + Noun: *movie theatre* vs *film theatre*
  - Adverb + Verb: *thoroughly amuse* vs *completely amuse*

Collocations

- Computational approaches generally make use of distributional differences for detecting and correcting errors
- Same general approach as in articles and prepositions:
  - Choose preferred form from a set of alternatives
  - But: the confusion set is potentially much larger
- Solution:
  - Constrain the space by selecting alternatives with a similar meaning
- See work on automatic thesaurus construction [eg Lin 1998]

Verb Form Errors

Error Type	Example
Subject-Verb Agreement	He have been living here since June.
Auxiliary Agreement	He has been live here since June.
Complementation	He wants live here.

- **See Lee and Seneff [2008] for a method based on detecting specific irregularities in parse trees.**

Outline

- Background
- Article Errors
- Preposition Errors
- Other ESL Problems
- Conclusions

Conclusions

- The provision of assistance to ESL learners is clearly a significant market
- Technology is at a very early stage, focussing on specific subproblems
- Measurable progress has been hampered by the unavailability of shared data sets, but this is changing