Correcting Misuse of Verb Forms

John Lee and Stephanie Seneff
Spoken Language Systems
MIT Computer Science and Artificial Intelligence Laboratory
Cambridge, MA 02139, USA
{jsylee,seneff}@csail.mit.edu

Abstract

This paper proposes a method to correct English verb form errors made by non-native speakers. A basic approach is template matching on parse trees. The proposed method improves on this approach in two ways. To improve recall, irregularities in parse trees caused by verb form errors are taken into account; to improve precision, n-gram counts are utilized to filter proposed corrections. Evaluation on non-native corpora, representing two genres and mother tongues, shows promising results.

1 Introduction

In order to describe the nuances of an action, a verb may be associated with various concepts such as tense, aspect, voice, mood, person and number. In some languages, such as Chinese, the verb itself is not inflected, and these concepts are expressed via other words in the sentence. In highly inflected languages, such as Turkish, many of these concepts are encoded in the inflection of the verb. In between these extremes, English uses a combination of inflections (see Table 1) and “helping words”, or auxiliaries, to form complex verb phrases.

It should come as no surprise, then, that the misuse of verb forms is a common error category for some non-native speakers of English. For example, in the Japanese Learners of English corpus (Izumi et al., 2003), errors related to verbs are among the most frequent categories. Table 2 shows some sentences with these errors.

Form	Example
base (bare)	*speak*
base (infinitive)	to *speak*
third person singular	*speaks*
past	*spoke*
-ing participle	*speaking*
-ed participle	*spoken*

Table 1: Five forms of inflections of English verbs (Quirk et al., 1985), illustrated with the verb “*speak*”. The base form is also used to construct the infinitive with “*to*”. An exception is the verb “*to be*”, which has more forms.

A system that automatically detects and corrects misused verb forms would be both an educational and practical tool for students of English. It may also potentially improve the performance of machine translation and natural language generation systems, especially when the source and target languages employ very different verb systems.

Research on automatic grammar correction has been conducted on a number of different parts-of-speech, such as articles (Knight and Chander, 1994) and prepositions (Chodorow et al., 2007). Errors in verb forms have been covered as part of larger systems such as (Heidorn, 2000), but we believe that their specific research challenges warrant more detailed examination.

We build on the basic approach of template-matching on parse trees in two ways. To improve recall, irregularities in parse trees caused by verb form errors are considered; to improve precision, n-gram counts are utilized to filter proposed corrections.

We start with a discussion on the scope of our

Proceedings of ACL-08: HLT, pages 174–182,
Columbus, Ohio, USA, June 2008.

task in the next section. We then analyze the specific research issues in §3 and survey previous work in §4. A description of our data follows. Finally, we present experimental results and conclude.

2 Background

An English verb can be inflected in five forms (see Table 1). Our goal is to correct confusions among these five forms, as well as the infinitive. These confusions can be viewed as symptoms of one of two main underlying categories of errors; roughly speaking, one category is semantic in nature, and the other, syntactic.

2.1 Semantic Errors

The first type of error is concerned with inappropriate choices of tense, aspect, voice, or mood. These may be considered errors in semantics. In the sentence below, the verb "*live*" is expressed in the simple present tense, rather than the perfect progressive:

*He *lives there since June.* (1)

Either "*has been living*" or "*had been living*" may be the valid correction, depending on the context. If there is no temporal expression, correction of tense and aspect would be even more challenging.

Similarly, correcting voice and mood often requires real-world knowledge. Suppose one wants to say "*I am prepared for the exam*", but writes "*I am preparing for the exam*". Semantic analysis of the context would be required to correct this kind of error, which will not be tackled in this paper[1].

[1]If the input is "*I am *prepare for the exam*", however, we will attempt to choose between the two possibilities.

Example	Usage
*I take a bath and *reading books.*	FINITE
*I can't *skiing well , but ...*	BASE_{md}
*Why did this *happened?*	BASE_{do}
*But I haven't *decide where to go.*	ED_{perf}
*I don't want *have a baby.*	INF_{verb}
*I have to save my money for *ski.*	ING_{prep}
*My son was very *satisfy with ...*	ED_{pass}
*I am always *talk to my father.*	ING_{prog}

Table 2: Sentences with verb form errors. The intended usages, shown on the right column, are defined in Table 3.

2.2 Syntactic Errors

The second type of error is the misuse of verb forms. Even if the intended tense, aspect, voice and mood are correct, the verb phrase may still be constructed erroneously. This type of error may be further subdivided as follows:

Subject-Verb Agreement The verb is not correctly inflected in number and person with respect to the subject. A common error is the confusion between the base form and the third person singular form, e.g.,

*He *have been living there since June.* (2)

Auxiliary Agreement In addition to the modal auxiliaries, other auxiliaries must be used when specifying the perfective or progressive aspect, or the passive voice. Their use results in a complex verb phrase, i.e., one that consists of two or more verb constituents. Mistakes arise when the main verb does not "agree" with the auxiliary. In the sentence below, the present perfect progressive tense ("*has been living*") is intended, but the main verb "*live*" is mistakenly left in the base form:

*He has been *live there since June.* (3)

In general, the auxiliaries can serve as a hint to the intended verb form, even as the auxiliaries "*has been*" in the above case suggest that the progressive aspect was intended.

Complementation A nonfinite clause can serve as complementation to a verb or to a preposition. In the former case, the verb form in the clause is typically an infinitive or an *-ing* participle; in the latter, it is usually an *-ing* participle. Here is an example of a wrong choice of verb form in complementation to a verb:

*He wants *live there.* (4)

In this sentence, "*live*", in its base form, should be modified to its infinitive form as a complementation to the verb "*wants*".

This paper focuses on correcting the above three error types: subject-verb agreement, auxiliary agreement, and complementation. Table 3 gives a complete list of verb form usages which will be covered.

Form	Usage	Description	Example
Base Form as Bare Infinitive	BASE_{md} BASE_{do}	After modals "Do"-support/-periphrasis; emphatic positive	He **may** *call*. **May** he *call*? He **did** not *call*. **Did** he *call*? I **did** *call*.
Base or 3rd person	FINITE	Simple present or past tense	He *calls*.
Base Form as *to*-Infinitive	INF_{verb}	Verb complementation	He **wants** her *to call*.
-ing participle	ING_{prog} ING_{verb} ING_{prep}	Progressive aspect Verb complementation Prepositional complementation	He **was** *calling*. **Was** he *calling*? He **hated** *calling*. The device is designed **for** *calling*
-ed participle	ED_{perf} ED_{pass}	Perfect aspect Passive voice	He **has** *called*. **Has** he *called*? He **was** *called*. **Was** he *called*?

Table 3: Usage of various verb forms. In the examples, the *italized* verbs are the "targets" for correction. In complementations, the main verbs or prepositions are **bolded**; in all other cases, the auxiliaries are **bolded**.

3 Research Issues

One strategy for correcting verb form errors is to identify the intended syntactic relationships between the verb in question and its neighbors. For subject-verb agreement, the subject of the verb is obviously crucial (e.g., "*he*" in (2)); the auxiliary is relevant for resolving auxiliary agreement (e.g., "*has been*" in (3)); determining the verb that receives the complementation is necessary for detecting any complementation errors (e.g., "*wants*" in (4)). Once these items are identified, most verb form errors may be corrected in a rather straightforward manner.

The success of this strategy, then, hinges on accurate identification of these items, for example, from parse trees. Ambiguities will need to be resolved, leading to two research issues (§3.2 and §3.3).

3.1 Ambiguities

The three so-called *primary verbs*, "*have*", "*do*" and "*be*", can serve as either main or auxiliary verbs. The verb "*be*" can be utilized as a main verb, but also as an auxiliary in the progressive aspect (ING_{prog} in Table 3) or the passive voice (ED_{pass}). The three examples below illustrate these possibilities:

This **is** *work not play.* (main verb)

My father **is** *working in the lab.* (ING_{prog})

A solution **is** *worked out.* (ED_{pass})

These different roles clearly affect the forms required for the verbs (if any) that follow. Disambiguation among these roles is usually straightforward because of the different verb forms (e.g., "*working*" vs. "*worked*"). If the verb forms are incorrect, disambiguation is made more difficult:

This **is** *work not play.*

My father **is** **work in the lab.*

A solution **is** **work out.*

Similar ambiguities are introduced by the other primary verbs[2]. The verb "*have*" can function as an auxiliary in the perfect aspect (ED_{perf}) as well as a main verb. The versatile "*do*" can serve as "do"-support or add emphasis (BASE_{do}), or simply act as a main verb.

3.2 Automatic Parsing

The ambiguities discussed above may be expected to cause degradation in automatic parsing performance. In other words, sentences containing verb form errors are more likely to yield an "incorrect" parse tree, sometimes with significant differences. For example, the sentence "*My father is *work in the laboratory*" is parsed (Collins, 1997) as:

```
(S (NP My father)
 (VP is (NP work))
 (PP in the laboratory))
```

[2]The abbreviations *'s* (*is* or *has*) and *'d* (*would* or *had*) compound the ambiguities.

The progressive form "*working*" is substituted with its bare form, which happens to be also a noun. The parser, not unreasonably, identifies "*work*" as a noun. Correcting the *verb* form error in this sentence, then, necessitates considering the *noun* that is apparently a copular complementation.

Anecdotal observations like this suggest that one cannot use parser output naively[3]. We will show that some of the irregularities caused by verb form errors are consistent and can be taken into account.

One goal of this paper is to recognize irregularities in parse trees caused by verb form errors, in order to increase recall.

3.3 Overgeneralization

One potential consequence of allowing for irregularities in parse tree patterns is overgeneralization. For example, to allow for the "parse error" in §3.2 and to retrieve the word "*work*", every determiner-less noun would potentially be turned into an *-ing* participle. This would clearly result in many invalid corrections. We propose using n-gram counts as a filter to counter this kind of overgeneralization.

A second goal is to show that n-gram counts can effectively serve as a filter, in order to increase precision.

4 Previous Research

This section discusses previous research on processing verb form errors, and contrasts verb form errors with those of the other parts-of-speech.

4.1 Verb Forms

Detection and correction of grammatical errors, including verb forms, have been explored in various applications. Hand-crafted error production rules (or "mal-rules"), augmenting a context-free grammar, are designed for a writing tutor aimed at deaf students (Michaud et al., 2000). Similar strategies with parse trees are pursued in (Bender et al., 2004), and error templates are utilized in (Heidorn, 2000) for a word processor. Carefully hand-crafted rules, when used alone, tend to yield high precision; they may, however, be less equipped to detect verb form errors within a perfectly grammatical sentence, such as the example given in §3.2.

An approach combining a hand-crafted context-free grammar and stochastic probabilities is pursued in (Lee and Seneff, 2006), but it is designed for a restricted domain only. A maximum entropy model, using lexical and POS features, is trained in (Izumi et al., 2003) to recognize a variety of errors. It achieves 55% precision and 23% recall overall, on evaluation data that partially overlap with those of the present paper. Unfortunately, results on verb form errors are not reported separately, and comparison with our approach is therefore impossible.

4.2 Other Parts-of-speech

Automatic error detection has been performed on other parts-of-speech, e.g., articles (Knight and Chander, 1994) and prepositions (Chodorow et al., 2007). The research issues with these parts-of-speech, however, are quite distinct. Relative to verb forms, errors in these categories do not "disturb" the parse tree as much. The process of feature extraction is thus relatively simple.

5 Data

5.1 Development Data

To investigate irregularities in parse tree patterns (see §3.2), we utilized the AQUAINT Corpus of English News Text. After parsing the corpus (Collins, 1997), we artificially introduced verb form errors into these sentences, and observed the resulting "disturbances" to the parse trees.

For disambiguation with n-grams (see §3.3), we made use of the WEB 1T 5-GRAM corpus. Prepared by Google Inc., it contains English n-grams, up to 5-grams, with their observed frequency counts from a large number of web pages.

5.2 Evaluation Data

Two corpora were used for evaluation. They were selected to represent two different genres, and two different mother tongues.

JLE (Japanese Learners of English corpus) This corpus is based on interviews for the Standard Speaking Test, an English-language proficiency test conducted in Japan (Izumi et al.,

[3]According to a study on parsing ungrammatical sentences (Foster, 2007), subject-verb and determiner-noun agreement errors can lower the F-score of a state-of-the-art probabilistic parser by 1.4%, and context-sensitive spelling errors (not verbs specifically), by 6%.

Input	Hypothesized Correction		
	None	Valid	Invalid
w/ errors	$false\_neg$	$true\_pos$	inv_pos
w/o errors	$true\_neg$	$false\_pos$	

Table 4: Possible outcomes of a hypothesized correction.

2003). For 167 of the transcribed interviews, totalling 15,637 sentences[4], grammatical errors were annotated and their corrections provided. By retaining the verb form errors[5], but correcting all other error types, we generated a test set in which 477 sentences (3.1%) contain subject-verb agreement errors, and 238 (1.5%) contain auxiliary agreement and complementation errors.

HKUST This corpus[6] of short essays was collected from students, all native Chinese speakers, at the Hong Kong University of Science and Technology. It contains a total of 2556 sentences. They tend to be longer and have more complex structures than their counterparts in the JLE. Corrections are not provided; however, part-of-speech tags are given for the original words, and for the intended (but unwritten) corrections. Implications on our evaluation procedure are discussed in §5.4.

5.3 Evaluation Metric

For each verb in the input sentence, a change in verb form may be hypothesized. There are five possible outcomes for this hypothesis, as enumerated in Table 4. To penalize "false alarms", a strict definition is used for false positives — even when the hypothesized correction yields a good sentence, it is still considered a false positive so long as the original sentence is acceptable.

It can sometimes be difficult to determine which words should be considered verbs, as they are not clearly demarcated in our evaluation corpora. We will thus apply the outcomes in Table 4 at the sentence level; that is, the output sentence is considered a true positive only if the original sentence contains errors, and only if valid corrections are offered for *all* errors.

The following statistics are computed:

Accuracy The proportion of sentences which, after being treated by the system, have correct verb forms. That is, ($true\_neg + true\_pos$) divided by the total number of sentences.

Recall Out of all sentences with verb form errors, the percentage whose errors have been successfully corrected by the system. That is, $true\_pos$ divided by ($true\_pos + false\_neg + inv\_pos$).

Detection Precision This is the first of two types of precision to be reported, and is defined as follows: Out of all sentences for which the system has hypothesized corrections, the percentage that actually contain errors, without regard to the validity of the corrections. That is, ($true\_pos + inv\_pos$) divided by ($true\_pos + inv\_pos + false\_pos$).

Correction Precision This is the more stringent type of precision. In addition to successfully determining that a correction is needed, the system must offer a valid correction. Formally, it is $true\_pos$ divided by ($true\_pos + false\_pos + inv\_pos$).

5.4 Evaluation Procedure

For the JLE corpus, all figures above will be reported. The HKUST corpus, however, will not be evaluated on subject-verb agreement, since a sizable number of these errors are induced by other changes in the sentence[7].

Furthermore, the HKUST corpus will require manual evaluation, since the corrections are not annotated. Two native speakers of English were given the edited sentences, as well as the original input. For each pair, they were asked to select one of four statements: one of the two is better, or both are equally correct, or both are equally incorrect. The

[4]Obtained by segmenting (Reynar and Ratnaparkhi, 1997) the interviewee turns, and discarding sentences with only one word. The HKUST corpus was processed likewise.

[5]Specifically, those tagged with the "v_fml", "v_fin" (covering auxiliary agreement and complementation) and "v_agr" (subject-verb agreement) types; those with semantic errors (see §2.1), i.e. "v_tns" (tense), are excluded.

[6]Provided by Prof. John Milton, personal communication.

[7]e.g., the subject of the verb needs to be changed from singular to plural.

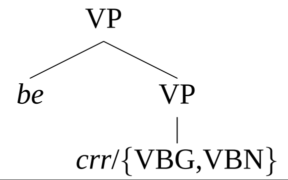
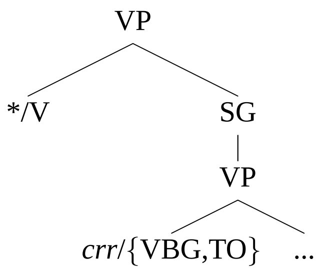
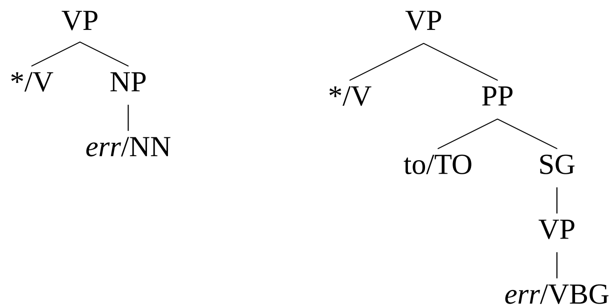
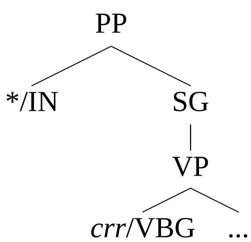
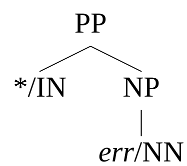
Expected Tree $\{\langle$*usage*$\rangle$,...$\}$	Tree disturbed by substitution $[\langle$*crr*$\rangle \rightarrow \langle$*err*$\rangle]$
$\{\text{ING}_{prog},\text{ED}_{pass}\}$	*A dog is [sleeping→sleep]. I'm [living→live] in XXX city.*
	
$\{\text{ING}_{verb},\text{INF}_{verb}\}$	*I like [skiing→ski] very much; She likes to [go→going] around*
	
ING_{prep}	*I lived in France for [studying→study] French language.*
	

Table 5: Effects of incorrect verb forms on parse trees. The left column shows trees normally expected for the indicated usages (see Table 3). The right column shows the resulting trees when the correct verb form ⟨*crr*⟩ is replaced by ⟨*err*⟩. Detailed comments are provided in §6.1.

correction precision is thus the proportion of pairs where the edited sentence is deemed better. Accuracy and recall cannot be computed, since it was impossible to distinguish syntactic errors from semantic ones (see §2).

5.5 Baselines

Since the vast majority of verbs are in their correct forms, the *majority baseline* is to propose no correction. Although trivial, it is a surprisingly strong baseline, achieving more than 98% for auxiliary agreement and complementation in JLE, and just shy of 97% for subject-verb agreement.

For auxiliary agreement and complementation, the *verb-only baseline* is also reported. It attempts corrections only when the word in question is actually tagged as a verb. That is, it ignores the spurious noun- and adjectival phrases in the parse tree discussed in §3.2, and relies only on the output of the part-of-speech tagger.

6 Experiments

Corresponding to the issues discussed in §3.2 and §3.3, our experiment consists of two main steps.

6.1 Derivation of Tree Patterns

Based on (Quirk et al., 1985), we observed tree patterns for a set of verb form usages, as summarized in Table 3. Using these patterns, we introduced verb form errors into AQUAINT, then re-parsed the corpus (Collins, 1997), and compiled the changes in the "disturbed" trees into a catalog.

N-gram	Example
`be` $\{$ING$_{prog}$, ED$_{pass}\}$ $*$	The dog `is` *sleeping*. The door `is` *open*.
`verb` $\{$ING$_{verb}$, INF$_{verb}\}$ $*$	I `need` *to do* this. I `need` *beef* for the curry.
$verb_1$ *ing `and` $\{$ING$_{verb}$, INF$_{verb}\}$	*enjoy* reading `and` *going* to pachinko *go* shopping `and` *have* dinner
`prep` $\{$ING$_{prep}\}$ $*$	`for` *studying* French language a class `for` *sign* language
`have` $\{$ED$_{perf}\}$ *	I `have` *rented* a video I `have` *lunch* in Ginza

Table 6: The n-grams used for filtering, with examples of sentences which they are intended to differentiate. The hypothesized usages (shown in the curly brackets) as well as the original verb form, are considered. For example, the first sentence is originally "*The dog is *sleep.*" The three trigrams "*is sleeping .*", "*is slept .*" and "*is sleep .*" are compared; the first trigram has the highest count, and the correction "*sleeping*" is therefore applied.

A portion of this catalog[8] is shown in Table 5. Comments on $\{$ING$_{prog}$,ED$_{pass}\}$ can be found in §3.2. Two cases are shown for $\{$ING$_{verb}$,INF$_{verb}\}$. In the first case, an *-ing* participle in verb complementation is reduced to its base form, resulting in a noun phrase. In the second, an infinitive is constructed with the *-ing* participle rather than the base form, causing "*to*" to be misconstrued as a preposition. Finally, in ING$_{prep}$, an *-ing* participle in preposition complementation is reduced to its base form, and is subsumed in a noun phrase.

6.2 Disambiguation with N-grams

The tree patterns derived from the previous step may be considered as the "necessary" conditions for proposing a change in verb forms. They are not "sufficient", however, since they tend to be overly general. Indiscriminate application of these patterns on AQUAINT would result in false positives for 46.4% of the sentences.

For those categories with a high rate of false positives (all except BASE$_{md}$, BASE$_{do}$ and FINITE), we utilized n-grams as filters, allowing a correction only when its n-gram count in the WEB 1T 5-GRAM corpus is greater than that of the original. The filtering step reduced false positives from 46.4% to less than 1%. Table 6 shows the n-grams, and Table 7 provides a breakdown of false positives in AQUAINT after n-gram filtering.

[8]Due to space constraints, only those trees with significant changes above the leaf level are shown.

Hyp. Usage	False Pos.	Hypothesized Usage	False Pos.
BASE$_{md}$	16.2%	$\{$ING$_{verb}$,INF$_{verb}\}$	33.9%
BASE$_{do}$	0.9%	$\{$ING$_{prog}$,ED$_{pass}\}$	21.0%
FINITE	12.8%	ING$_{prep}$	13.7%
		ED$_{perf}$	1.4%

Table 7: The distribution of false positives in AQUAINT. The total number of false positives is 994, represents less than 1% of the 100,000 sentences drawn from the corpus.

6.3 Results for Subject-Verb Agreement

In JLE, the accuracy of subject-verb agreement error correction is 98.93%. Compared to the majority baseline of 96.95%, the improvement is statistically significant[9]. Recall is 80.92%; detection precision is 83.93%, and correction precision is 81.61%.

Most mistakes are caused by misidentified subjects. Some *wh*-questions prove to be especially difficult, perhaps due to their relative infrequency in newswire texts, on which the parser is trained. One example is the question "*How much extra time does the local train *takes?*". The word "*does*" is not recognized as a "do"-support, and so the verb "*take*" was mistakenly turned into a third person form to agree with "*train*".

6.4 Results for Auxiliary Agreement & Complementation

Table 8 summarizes the results for auxiliary agreement and complementation, and Table 2 shows some examples of real sentences corrected by the system. Our proposed method yields 98.94% accuracy. It is a statistically significant improvement over the majority baseline (98.47%), although not significant over the verb-only baseline[10] (98.85%), perhaps a reflection of the small number of test sentences with verb form errors. The Kappa statistic for the man-

[9]$p < 0.005$ according to McNemar's test.

[10]With $p = 1 * 10^{-10}$ and $p = 0.038$, respectively, according to McNemar's test

Corpus	Method	Accuracy	Precision (correction)	Precision (detection)	Recall
JLE	verb-only	98.85%	71.43%	84.75%	31.51%
	all	98.94%	68.00%	80.67%	42.86%
HKUST	all	not available	71.71%	not available	

Table 8: Results on the JLE and HKUST corpora for auxiliary agreement and complementation. The majority baseline accuracy is 98.47% for JLE. The verb-only baseline accuracy is 98.85%, as indicated on the second row. "All" denotes the complete proposed method. See §6.4 for detailed comments.

Usage	JLE	HKUST
	Count (Prec.)	Count (Prec.)
BASE_{md}	13 (92.3%)	25 (80.0%)
BASE_{do}	5 (100%)	0
FINITE	9 (55.6%)	0
ED_{perf}	11 (90.9%)	3 (66.7%)
$\{\text{ING}_{prog},\text{ED}_{pass}\}$	54 (58.6%)	30 (70.0%)
$\{\text{ING}_{verb},\text{INF}_{verb}\}$	45 (60.0%)	16 (59.4%)
ING_{prep}	10 (60.0%)	2 (100%)

Table 9: Correction precision of individual correction patterns (see Table 5) on the JLE and HKUST corpus.

ual evaluation of HKUST is 0.76, corresponding to "substantial agreement" between the two evaluators (Landis and Koch, 1977). The correction precisions for the JLE and HKUST corpora are comparable.

Our analysis will focus on $\{\text{ING}_{prog},\text{ED}_{pass}\}$ and $\{\text{ING}_{verb},\text{INF}_{verb}\}$, two categories with relatively numerous correction attempts and low precisions, as shown in Table 9. For $\{\text{ING}_{prog},\text{ED}_{pass}\}$, many invalid corrections are due to wrong predictions of voice, which involve semantic choices (see §2.1). For example, the sentence "*... the main duty is study well*" is edited to "*... the main duty is studied well*", a grammatical sentence but semantically unlikely.

For $\{\text{ING}_{verb},\text{INF}_{verb}\}$, a substantial portion of the false positives are valid, but unnecessary, corrections. For example, there is no need to turn "*I like cooking*" into "*I like to cook*", as the original is perfectly acceptable. Some kind of confidence measure on the n-gram counts might be appropriate for reducing such false alarms.

Characteristics of speech transcripts pose some further problems. First, colloquial expressions, such as the word "*like*", can be tricky to process. In the question "*Can you like give me the money back*", "*like*" is misconstrued to be the main verb, and "*give*" is turned into an infinitive, resulting in "*Can you like *to give me the money back*". Second, there are quite a few incomplete sentences that lack subjects for the verbs. No correction is attempted on them.

Also left uncorrected are misused forms in non-finite clauses that describe a noun. These are typically base forms that should be replaced with *-ing* participles, as in "*The girl *wear a purple skiwear is a student of this ski school*". Efforts to detect this kind of error had resulted in a large number of false alarms.

Recall is further affected by cases where a verb is separated from its auxiliary or main verb by many words, often with conjunctions and other verbs in between. One example is the sentence "*I used to climb up the orange trees and *catching insects*". The word "*catching*" should be an infinitive complementing "*used*", but is placed within a noun phrase together with "*trees*" and "*insects*".

7 Conclusion

We have presented a method for correcting verb form errors. We investigated the ways in which verb form errors affect parse trees. When allowed for, these unusual tree patterns can expand correction coverage, but also tend to result in overgeneration of hypothesized corrections. N-grams have been shown to be an effective filter for this problem.

8 Acknowledgments

We thank Prof. John Milton for the HKUST corpus, Tom Lee and Ken Schutte for their assistance with the evaluation, and the anonymous reviewers for their helpful feedback.

References

E. Bender, D. Flickinger, S. Oepen, A. Walsh, and T. Baldwin. 2004. Arboretum: Using a Precision Grammar for Grammar Checking in CALL. *Proc. InSTIL/ICALL Symposium on Computer Assisted Learning*.

M. Chodorow, J. R. Tetreault, and N.-R. Han. 2007. Detection of Grammatical Errors Involving Prepositions. In *Proc. ACL-SIGSEM Workshop on Prepositions*. Prague, Czech Republic.

M. Collins. 1997. Three Generative, Lexicalised Models for Statistical Parsing. *Proc. ACL*.

J. Foster. 2007. Treebanks Gone Bad: Generating a Treebank of Ungrammatical English. In *Proc. IJCAI Workshop on Analytics for Noisy Unstructured Data*. Hyderabad, India.

G. Heidorn. 2000. Intelligent Writing Assistance. *Handbook of Natural Language Processing*. Robert Dale, Hermann Moisi and Harold Somers (ed.). Marcel Dekker, Inc.

E. Izumi, K. Uchimoto, T. Saiga, T. Supnithi, and H. Isahara. 2003. Automatic Error Detection in the Japanese Learner's English Spoken Data. In *Companion Volume to Proc. ACL*. Sapporo, Japan.

K. Knight and I. Chander. 1994. Automated Postediting of Documents. In *Proc. AAAI*. Seattle, WA.

J. R. Landis and G. G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33(1):159–174.

L. Michaud, K. McCoy and C. Pennington. 2000. An Intelligent Tutoring System for Deaf Learners of Written English. *Proc. 4th International ACM Conference on Assistive Technologies*.

J. Lee and S. Seneff. 2006. Automatic Grammar Correction for Second-Language Learners. In *Proc. Interspeech*. Pittsburgh, PA.

J. C. Reynar and A. Ratnaparkhi. 1997. A Maximum Entropy Approach to Identifying Sentence Boundaries. In *Proc. 5th Conference on Applied Natural Language Processing*. Washington, D.C.

R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, New York.