文法誤り訂正における問題点の考察と新タスクの提案

永田　亮†
† 甲南大学知能情報学部
E-mail: †nagata-nlp@ hyogo-u.ac.jp.

1. はじめに

自由記述英作文を対象とした，語学学習支援のための文法誤り訂正は様々な手法が提案されており，訂正性能の改善は目覚ましい．冠詞，前置詞など幅広い項目について，専用の手法が考案されている．最近では，複数種類の誤りを同時に訂正することで更なる訂正性能の改善が達成されている．

一方で，学習効果という観点からは大きな問題が残されている．詳細は 2. に譲るが，その問題は次のように要約される：「現状の文法誤り訂正技術で得られる質と量の情報では学習効果が非常に低い」．語学学習の第一義的な目的は語学能力の向上であるから，学習効果が低いというのは重大である．著者が知る限り，そもそも，このような問題が存在することを明示的に議論した文献は存在しない．

このような背景を踏まえ，本稿では，学習効果を高めることを目的とした新しいタスクを提案する．まず，2. で，論理的考察により上述の問題を明らかにする．その結果を受けて，3. では，より高い学習効果が期待できる新タスク**内文法状態推定**を提案する．内文法状態推定とは，対象としている文法項目に関して，学習者がどの部分に問題を抱えているのかを検知するタスクである．誤り箇所と訂正候補を特定する文法誤り訂正と異なり，内文法状態推定は学習者が抱える問題自体を特定する．そのため，内文法状態推定では，なぜ誤りであるかを学習者に直接説明できるという利点がある．

本研究の貢献は，新しいタスクの提案に加え，次の 3 点である：(1) 人手で作成した比較的単純なルールで，前置詞誤りの理由をある程度特定できることを示す；(2) 特定結果を基に内文法状態を推定するための方法論を示す；(3) その有効性を評価し，推定精度を更に向上させるためのアイデアを示す．なお，本項では，主に，前置詞誤りを対象として議論を進めることにする．

2. 文法誤り訂正における問題点の考察

結論から述べると，ある種の文法誤りにおいて現状の誤り訂正技術が学習者に与える学習効果は非常に低い．このことを考察してみよう．

いま，語学学習の一環として，ある一定期間と回数（例えば半年 10 回），毎週 200 トークン程度からなるエッセイを学習者に書かせることを考えよう．このとき，文法誤り訂正技術で訂正できる前置詞誤りは非常に少ない．例えば，日本人英語学習者のエッセイを収録した Konan-JIEM Learenr Corpus [1]（以下，KJ コーパスと省略）では，前置詞誤りの割合はトークンあたり約 1.6%であるので，この値を参考にすると毎回 3 個程度の前置詞誤りが出現することになる．一方で，同コーパスを対象とした場合，現状で訂正性能が最も高い手法 [2] でも，訂正率 24%，訂正精度 37%である．したがって，平均的には，正しい訂正は毎回 1 個にも満たない（正確には，3 個× 24%=0.7 個程度）．また，正しい訂正より誤訂正のほうが多い傾向となる．更に悪いことに現在主流である分類器に基づく訂正手法では，なぜ誤りであるかを学習者に説明することができない．このような質と量のフィードバックでは，半年 10 回の学習を経ても，学習効果は非常に薄いであろう．上述の誤り率や訂正性能は KJ コーパス特有のものではなく，他のコーパスでも同様な値を示すことから，前置詞誤り訂正に共通する問題といえる．

訂正性能がそれほど高くなく，文法規則が明確でない誤り（例えば，時制・相の誤り）でも同様な傾向となると予想される(注 1)．文法誤り訂正は盛んに研究されているにも関わらず訂正性能がそれほど高くないことを考慮すると，短期的に訂正性能が大幅に改善される見込みは低いであろう．そうであれば，学習効果が低いという問題も短期的に解決される見込みは低いといえる．

語学学習の第一義的目的は語学能力の向上であるから，学習効果が低いというのは学習者を対象とした支援としては大きな問題である．訂正性能の改善という直接的な解決が難しい現状を踏まえると，何らかの別の方法でこの問題を解決することが望ましい．

3. 新タスク：内文法状態推定

2. で指摘した問題を解決する手段として，新しいタスク**内文法状態推定**を提案する．最初に，内文法状態推定を定義しておこう．まず，ある文法項目における誤りが，誤り理由に基づいていくつかのカテゴリに分類できると仮定する．例えば，前置詞の用法であれば，カテゴリは「前置詞は基本的に名詞句と共に使用されることを理解していない誤り」などになる．内文法状態とは，このカテゴリのうち，どれに学習者が問題を抱えているかを表すものとする．具体的には，学習者が各カテゴリに問題を抱えている／いないを，1 ／ 0 で表

(注 1)：一方で，主語–動詞の一致の誤りのように，比較的文法規則が明確で，訂正性能が高い誤りについては，現状の文法誤り訂正手法でも学習効果があるであろう．定期的に（たとえ，数回に一回でも）主語–動詞の一致の誤りを正しく訂正できれば，学習者は主語–動詞の一致を意識するようになるであろう．

す 2 値ベクトルとして表現する（もしくは，図 1 下部に示すように，連続値を仮定しても良い）．したがって，内文法状態推定は，ベクトルの値を決定する問題として定義できる．内文法状態推定の厳密な定義は以上の通りであるが，このままでは（おそらく文法誤り訂正より）難しいタスクである．

そこで，内文法状態推定を近似的に解くことを考える．状態を直接推定（すなわちベクトルの値を推定）する代わりに，各カテゴリの誤りが発生する確率を推定する問題に近似する．誤り確率が高精度で推定できれば，内文法状態の良い近似となるであろう．すなわち，誤り確率が高いカテゴリは，内文法状態の真の値が 1 である可能性も高い．

このように近似された内文法状態でも，学習効果を高めるという目的に十分有益である．内文法状態から問題を抱えるカテゴリがわかるので，例えば「前置詞は基本的に名詞と共に使用されます．名詞以外と前置詞を使用していないかチェックしてみましょう.」のように，直接的なフィードバックを学習者に与えることができる．また，教師にアラートを出すという利用方法もある．このようなフィードバックやアラートは，従来の文法誤り訂正では実現が困難である．

内文法状態推定が有効となるためには，少なくとも次の二つの疑問に答えなければならない：

（1） 誤り理由により，ある文法項目に関する誤りを細分類することができるのか？

（2） 誤り確率を高い精度で推定できるのか？

（1）と（2）について，それぞれ 4. と 5. で詳細に論ずる．

4. 前置詞誤り理由の分析

前置詞の用法を対象として誤り理由を分類するため，文法書と学習者コーパスを分析した．学習者コーパスについては，独自に収集した日本人英語学習者（中学生と高校生）の英文（2,140 文，19,368 トークン，前置詞誤り数 251）を用いた．

表 1 に分類結果を示す．同表に示すとおり，少なくとも九つのカテゴリが確認できた．以下，各カテゴリを詳細に説明するが，カテゴリのラベルと意味の対応については，表 1 の第一カラムを参照されたい．なお，上述のコーパスには出現しなかったが，理論上可能な誤り理由も表 1 に含めた．

カテゴリ PNN は，前置詞が副詞句など名詞句以外の語句と使用された誤りである（例：“He went *to there.”）．前置詞は名詞相当句と共に使用するという基本規則が理解できていない誤りである(注 2)．F も前置詞の基本的な用法である．主語と目的語以外の名詞句は，基本的に何らかの前置詞を伴わないといけないが，そのことが理解されていない誤りである（例：“He went there *this summer vacation.”）．

カテゴリ NWP は F の例外にあたる．すぐ上で，主語と目的語以外の名詞句は何らかの前置詞を伴うことが普通であることを述べたが，“today” や “this morning” など一部の名詞句では前置詞を必要としない．そのため，“He went there *in this morning.” などの誤りが起こる．

表 1： 前置詞誤り理由の分類

カテゴリ	頻度	割合 (%)
PNN：名詞句以外と使用された前置詞	6	2.4
F：浮いた名詞句	12	4.8
NWP：前置詞を必要としない名詞句	3	1.2
MT：to が抜けている不定詞	11	4.4
T：to 不定詞と動名詞の使い分け	1	0.4
ET：to が余分な不定詞	0	0.0
DP：直接目的語と前置詞句の使い分け	83	33.1
V：valence 誤り	7	2.8
S：選択誤り	111	44.2
その他	17	6.8
TOTAL	251	100.0

To-不定詞に関する誤りとして，カテゴリ MT， ET， T の 3 種類がある．MT は，ここでは，動詞が連続した誤り（例：“He wants *sing.”）を対象とした．動詞の連続は，一部の言語では許される（例：“Il veut chanter.”）ため，母語干渉として同種の誤りが起こると予想できる．ET は，逆に，原形不定詞が要求される文脈で，“to” を使用した誤りである（例：“He can to sing.”）．T は，to-不定詞の代わりに動名詞を取る動詞で to-不定詞を使用した誤りである（例：“He kept *to push.”）．

カテゴリ DP と V は，動詞の項に関する誤りである．前者は，前置詞句でなく直接目的語を項として取った誤り（例：“He went *Tokyo.”）とその逆の誤り（例：“He visited to Tokyo.”）である．また，valence とは動詞がいくつの項を取りうるかということを表す．例えば， 4 文型を取ることができる動詞は，主語，直接目的語，間接目的語の 3 つの項を取るため valence の値は 3 となる．Valence の概念を理解することは前置詞を正しく使用する上で重要である．例えば，valence を考慮しないと “He *explained me the detail” などの誤りが起こる．

カテゴリ S は，抜けと余剰以外の選択誤りである．この誤りは前置詞および関連する名詞句と動詞句により細分類することが可能であるが，個別の事例が多く分類が多岐に渡ることから，ここでは暫定的に一つのカテゴリにまとめた．

以上より，3. で導入した疑問（1）に対して，「少なくとも前置詞の誤りに関しては，誤り理由により分類可能である」と答えることができる．同様に，冠詞の用法（例：「冠詞は名詞句内で使用する」，「名詞の可算性に関する用法」，「定／不定の用法」）などでも分類可能であろう．

5. パイロットスタディ：内文法状態の推定

5.1 内文法推定手法

推定手法の概要は次の通りである（図 1 も参照のこと）．

(注 2)：もしくは，品詞の概念が理解できていない可能性もある．いずれにせよ，このカテゴリが習得できてないことが検知できれば，品詞と前置詞の基本的用法を学習者に説明すれば良い．

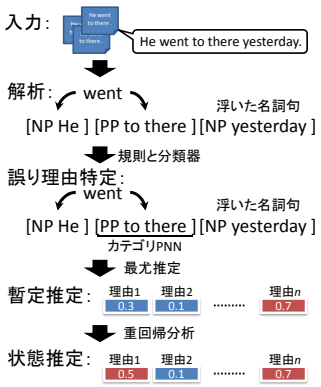

図 1： 内文法状態推定の流れ

入力は，内文法状態推定の対象となる学習者が書いた複数のエッセイとする．まず，入力エッセイに品詞解析と句解析を適用して前置詞と動詞／名詞句の関係を決定する．正確に関係を把握するためには構文解析が必要となるが，本稿では，品詞解析と句解析のみを利用する．なぜなら，対象が学習者の英文であり，構文解析の精度が低い恐れがあるためである(注 3)．次に，人手で作成した規則と分類器を解析結果に適用して，誤り理由の特定を行う．その結果を基に，各誤り理由の出現確率を推定し暫定値とする．最後に，暫定値に重回帰分析を適用して誤り確率を再推定する．推定結果を集めたものが内文法状態推定の結果となる．以降，各処理の詳細について述べる．なお，以下の規則や閾値は，4. で述べた学習者コーパスに基づいて決定した．

前置詞と動詞／名詞句との関係の解析は単純な規則に基づいて行う．品詞解析と句解析の結果から，前置詞の右隣の句を前置詞句とする．また，動詞と名詞句の関係も決定する．まず，代名詞のように語形から主語／目的語の判断がつく名詞句を対応する動詞に割り当てる．その際には，語順の制約（主語は動詞句の左側，目的語は右側）を満たす直近の動詞句に対応する名詞句を割り当てる．残りの名詞句も語順の制約を満たしながら動詞句に割り当てる．ただし，主語は最大で一つ，目的語については二つという制約も課す．更に，名詞句の探索は前置詞を越えては行わないという制約も課す．その結果，どの動詞にも割り当てられない（浮いた）名詞句も生じる（これは誤りである可能性が高い）．この時点で，各動詞の valence の値も決定できる．

以上の解析結果に基づいて誤り理由の特定を行う．基本は，人手で作成した規則で行う．各規則は次のとおりである．カテゴリ PNN は，前置詞句内に形容詞句，副詞句，主動詞を含む動詞句のいずれかであれば誤りとする．この規則（および以降，全ての規則）には，例外が生じることは容易に想像できるが，例外については別処理で対応することとする．カテゴリ F については，解析の結果，どの動詞句にも割り当てられない名詞句とする．更に，主動詞が二つ連続している場合をカテゴリ MT とする．逆に，to-不定詞を必要としない環境で “to” が使用されている場合を ET とする．具体的には，(i) 助動詞と主動詞の間，(ii) 原形不定詞を必要とする動詞の右隣二つの句とする．原形不定詞を必要とする動詞については，“feel”，“find”，“hear”，“let”，“make”，“notice”，“perceive”，“see”，“watch” の 9 種類とする．ET についても，文法書を参考にして人手で作成した名詞句辞書に基づいて検出する．登録された名詞句（例えば，“today” や “this morning”）が，“at”，“in”，“on”，“to” のいずれかと使用されている場合誤りとする．

残りのカテゴリについては，母語話者コーパスから自動獲得した辞書に基づいて検出する．上述の解析処理を母語話者コーパスに適用することで valence 辞書を得る．具体的には，ある閾値以上の用法をその動詞が取りうる valence の値とする（ここでは，頻度が 5 以上かつ割合が 10%以上のものを採択した）．この valence 辞書と解析の結果得られる valence の値を比較することでカテゴリ V の誤りを特定する．また，各動詞が直接目的語と前置詞句のどちら（あるいは両方）を取りうるかを決定する辞書も母語話者コーパスから得る．動詞句の右隣の句が名詞句か前置詞句かの頻度を求め辞書とする（この場合も上と同じ閾値を採択した）．同様に，to-不定詞と動名詞の使い分けに関する辞書も得る．これらの辞書を用いることで，カテゴリ DP と T に関する誤りを検出する．

既に述べたように，以上の手法は単純な規則に基づくため誤検出が頻出すると予想される．そこで例外処理を行う．まず，以上の手法を母語話者コーパスに適用する．その際，いずれかの誤り理由が特定されたとすると誤検出である可能性が高い．したがって，検出された表層パターンを例外的に正しい表現として例外表現辞書に登録する．表層パターンは，誤りが検出された位置から右隣二つの句内のトークンとする（ただし，カテゴリ F の場合のみ，名詞句内のトークンとする）．同様に，その二つの句を中心として前後 L 個（本稿では $L = 5$ とする）の句内から得られる品詞ラベルおよび句ラベルの連鎖の全ての組み合わせも例外表現の候補とする．母語話者コーパスにおける頻度および部分文字列との頻度比が閾値（本稿ではそれぞれ 5 と 0.5）以上のものを例外表現辞書に登録する．以上の処理から得られる例外表現に一致するものは，誤り特定処理の際に誤検出と判断し誤りとしない．

カテゴリ S については，誤り格フレームに基づいた手法 [2] と分類器に基づいた手法 [3] とを利用する．前者については，選択誤り以外にも誤り理由を特定できるため，その情報から上述の誤り理由の特定にも利用する．

最終的な誤り理由の特定は，全ての検出結果の論理和を取ることで行う．同じ箇所で複数の誤り理由が特定された場合，規則に基づく手法，誤り格フレームに基づいた手法，分類器に基づいた手法の順で優先順位が高いとする．

以上の処理結果から内文法状態が一応のところ求まる．検

(注 3)：品詞解析と句解析については，学習者コーパスに対しても一定の精度が保てることが知られている [1].

出された誤り理由の頻度を（本稿では，名詞句の数で）正規化することで誤り確率を推定できる．この値を誤り確率の暫定値とする．

更に，暫定値の推定精度を高めることを考える．具体的な手段として重回帰分析を用いる．すなわち，誤り確率の暫定値を説明変数として，誤り確率を予測する重回帰分析である．どの暫定値を説明変数とするかは変数増加法により決定する．また，切片の値は 0 とする．

5.2 内文法状態推定実験

対象データとして KJ コーパスを使用した．KJ コーパスには，25 人分の英文が収集されており，各人は最大で 10 個のトピックについて英文を記述している（平均で 9.3 個）．本パイロットスタディでは，一人の学習者が書いた複数の英文を一つにまとめたものを一つの事例とした．言い換えれば，平均 9.3 個の英文から内文法状態を推定することになる．KJ コーパスで規定されている標準的な分割に従い，17 人分を訓練データ，残りを評価データとした．また，4. で用いた学習者コーパスも開発データとして使用した．母語話者コーパスとしては，英語教材からなる英文（16,010 文，214,622 トークン）を使用した．

表 2 に，評価データに対する誤り理由の特定結果を示す．表から，単純な規則でも，ある程度の性能で特定できるカテゴリがあることが分かる．例えば，基本的な用法である PNN については，7 割程度の誤りを精度 5 割で検出できている．一方で，F は全く検出できていない（開発データでも同じ傾向を示した）．なお，DP と V については，相互に関連が深く，開発データを用いた評価結果から，頻繁に混同して検出されることが明らかとなったため，表 2（および以降）では両者を区別せずまとめて V と表記している．

誤り特定処理の結果を利用して，内文法状態推定を行った．本稿では，紙面の関係から，特に誤り特定性能が低い F と S についてのみ，内文法状態推定の結果を報告する．なお，重回帰分析の結果，F については，V と NWP，V については S と V が説明変数として選択された．

評価尺度は，実際の誤り率と推定された誤り確率の誤差の二乗平均平方根とした（したがって，値が小さいほど性能が良い）．比較対象として，次の二つのベースラインを設定した．一つ目は誤り理由特定処理で得られた誤り確率の暫定値である．このベースラインよりも誤差が大きい場合，重回帰分析による誤り確率の推定の効果がないことを意味する．二つ目は，訓練データにおける実際の誤り率の平均値を推定値としたものである．

表 3 に結果を示す．表より，重回帰分析で誤り確率の推定精度が改善していることがわかる．言い換えれば，他の誤り理由の特定結果が，内文法状態推定に有益な情報をもたらすことを示唆する．一方で，訓練データの平均値に比べると，S では誤差が大きい．また，F では誤差が小さいが，両者の差は統計的には有意ではなかった．

表 2： 誤り理由の特定結果

カテゴリ	頻度	Recall	Precision	F_1
PNN	7	0.71	0.50	0.59
F	16	0.00	0.00	0.00
NWP	5	0.40	1.00	0.57
MT	6	0.50	0.75	0.60
ET	2	1.00	0.50	0.67
V	41	0.27	0.38	0.31
S	38	0.29	0.26	0.27
その他	14	0.00	0.00	0.00
TOTAL	129	0.26	0.36	0.30

表 3： 内文法状態の推定結果（誤り確率の推定誤差）

カテゴリ	誤り特定結果	訓練データ平均値	提案手法
F	1.30	0.96	0.77
S	0.98	0.55	0.78

以上の結果から，提案手法は内文法状態推定に一定の効果はあるが，更なる改善が必要であるといえる．改善策としては，まずデータ数を増やすことが考えられる．今回は，訓練データが 17 事例，評価データが 8 事例とデータ数が限られていた．この数を増やすことにより重回帰モデル，評価結果ともに信頼性を高められる．ただし，同一の学習者が書いた多数のエッセイ（今回は平均で 9.3 個）が必要であるところに難しさがある．現在公開されている学習者コーパスでは一人の学習者に対して 1〜数個のエッセイというのが主流である．また，推定モデルの改善方法としては，使用する情報を増やすということが挙げられる．例えば，言語モデルを使用すれば対象学習者の英文の自然さが定量化できるのでこの値を説明変数の一つにできる．更に，推定モデルを非線形なモデルにすることも必要かもしれない．

6. おわりに

本稿では，学習効果という観点からは，従来盛んに研究されてきた文法誤り訂正に大きな問題があることを指摘した．この問題の解決策として新しいタスク内文法状態推定を提案し，その解法の一例を示した．更に，実データを用いて次の 3 点を確認した：(1) 誤り理由により前置詞誤りを細分類できる；(2) 人手で作成した単純な規則で，ある程度誤り理由を特定できる；(3) 誤り理由の特定結果に基づいて，内文法状態の推定精度が改善できる．

参考文献

[1] R. Nagata et al., “Creating a manually error-tagged and shallow-parsed learner corpus,” Proc. of 49th ACL, pp.1210–1219, 2011.

[2] R. Nagata et al., “Correcting preposition errors in learner English using error case frames and feedback messages,” Proc. of 52nd ACL, pp.754–764, 2014.

[3] I. Yoshimoto et al., “NAIST at 2013 CoNLL grammatical error correction shared task,” Proc. of 17th CoNLL Shared Task, pp.26–33, 2013.