



教師の英作文添削を支援する誤り検出システム

川村, 晃市
柏木, 治美
康, 敏

(Citation)

日本教育工学会研究報告集, 19(1):83-88

(Issue Date)

2019

(Resource Type)

journal article

(Version)

Version of Record

(URL)

<https://hdl.handle.net/20.500.14094/90005737>



教師の英作文添削を支援する誤り検出システム

An Error Detection System for Teachers in Learners' English Writing

川村 晃市*

柏木 治美**

康 敏*

Koichi Kawamura*

Harumi Kashiwagi**

Min Kang*

神戸大学大学院国際文化学研究科* 神戸大学大学教育推進機構**

Graduate School of Intercultural Studies, Kobe University*

Institute for Promotion of Higher Education, Kobe University**

＜あらまし＞ 添削による教師の負担問題は英語ライティング授業における大きな課題と指摘されている。その課題の解決策として教育現場に英作文誤り検出システムの導入が考えられる。英作文誤りの自動検出手法に関する研究は自然言語処理技術の著しい発展に伴って盛んに行われ、多くのシステムは提案されているが、そのほとんどはローカルエラーを検出対象としている。そのためグローバルエラーの検出に焦点を当てたシステムはまだ少ない。そこで本研究では、学生が日本語エッセイから自由英作文を産出するライティング授業において、日英対照文型に基づいて構築した判定基準を用いて、英語の基本的構文となる主語と述語の誤り検出システムを構築した。

＜キーワード＞ 誤用検出 英語 日英対照文型 主語 述語

1. はじめに

英語学習者にとって、4 技能のうちライティングは「自然と習得されるものではなく、学校などで教育を受けなければ習得することが困難である」(Grabe and Kaplan 1996) ことから、他の 3 技能と比較して習得することが困難な技能であると言える。教育現場においては、ライティング活動の展開に多くの問題が指摘されている(木村ほか 2010)。中学・高校の英語教育に関しては、オーラル重視で、ライティングが後回しにされてしまう。また、大学の英語教育に関しても、4 技能に配分される時間のバランスには、理想と現実のギャップがあり、ライティングの時間が少ないという問題が存在している。内容に関しても、「文型や表現の練習」や「和文英訳」などがライティングの主流であり、エッセイなどの自由作文への指導は困難とされている。これらの問題に共通する原因として、言語習得上の特性はもちろんだが、添削の大変さと時間がかかることも挙げられる。ライティングの文法修正フィードバックの効果に否定的な見解(Truscott 1996)と肯定的な見解(Ferris 2004)が存在し、統一された指導方法が確立されていないことから、学習者の多様な誤用に対応しなければならず、指導する

にあたり準備に負担がかかる技能であると言える。

2004 年以降、大学などの高等教育機関では学習者のライティング能力向上を目的としたライティングセンターの設置が進められている(吉田 2010, 太田ほか 2013)が、まだ、その数は十分ではなく、利用できる回数および時間に制約があるため、学習者への支援は限定的なものに留まる。また、ライティングセンターの多くが「自立した書き手」を育成するために、支援の重点を思考プロセスおよび内容の推敲に置いている(吉田 2010)ことから、文単位の文章校正が行われないケースも少なくない。このようなライティングセンターの支援は正確な文法知識を有さない初級・中級学習者のライティング能力向上には適していない。これらの問題に対処するためには、さまざまな習熟度の学習者が独力でライティング能力を身につけることができる英作文支援システムの構築が必要であると同時に、教師の添削負担を軽減できる英作文添削支援システムの構築も不可欠であると考えられる。

英作文添削の中でも、重要なプロセスの一つとして誤用検出がある。学習者の誤用には様々なものが含まれるが、特に global error

は文の全体構造に影響して内容の理解を阻害するために、教師が英作文を添削する際には見落とすことは許されない。しかしながら、**global error**を検出するには誤用文の構造および学習者の意図を理解する必要があり、容易ではない。既存の誤り検出システムの多くは**local error**を対象としており、**global error**に焦点を当てたシステムは管見の限り存在しない。

そこで本研究では、教師の英作文添削を支援することを目標とし、学生が日本語エッセイから自由英作文を産出するライティング授業において、日英対照文型に基づいて構築した判定基準を用い、英語の基本的構文となる主語と述語の誤り検出システムを構築した。構築したシステムを利用して実際の学生の英作文を分析し、英作文の添削を支援できるポイントを明らかにした。

2. システムの概要

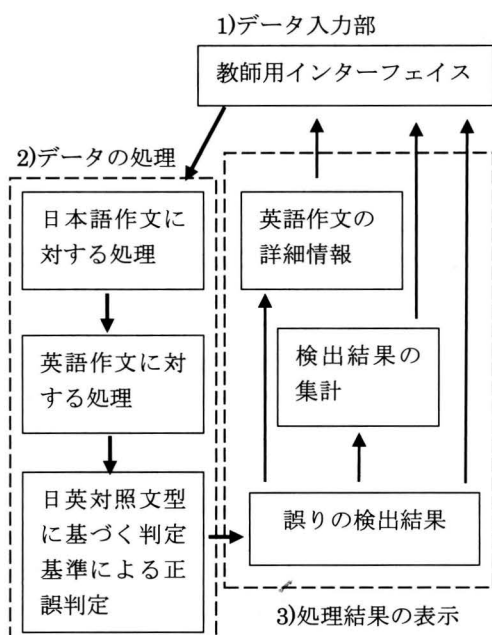


図1 システムの概要

本システムの目的は教師の英作文添削支援であるため、教師が実際に教育現場で使用する

ることを想定している。それゆえに、教師がすでに学習者の作文データ（日本語作文とそれに対応する英語作文）の有し、それらの作文データをシステムに入力できる段階にあることを前提としている。

システムは、教師が英語学習者の日本語作文およびそれに対応する英語作文を入力すると、規定した誤用判定基準にしたがって英語作文に含まれる誤りを検出し、その検出結果および集計結果を得ることができる仕様となっている。また、各英語作文の詳細情報も表示されるように設計されている。システムの構成は、1) データを入力する部分、2) そのデータを処理する部分、そして 3) 処理した結果を表示する部分からなっている。1) データ入力部分、および 3) 処理結果を表示する部分は PHP によって構築し、2) データを処理する部分については Python 3 を用いて構築している。図1はシステムの概要図である。

3. データを処理する部分

データの処理部分では、入力された日本語作文をセンテンスごとに主語と述語を抽出し、後述する日本語の文型に分類したうえで、そのセンテンスに対応する英語作文のセンテンスから主語と述語の抽出を試みる。そしてこれまでに構築した日英対照文型に基づく判定基準を用いて、誤りの抽出を行う。

3.1. 日本語及び英語の主語述語の抽出

データを処理するアルゴリズムは以下の例文を用いて説明する。

日本語文：「私は大学生でした。」

英語作文：“I am a college student.”

日本語のセンテンスの主語と述語の抽出には、日本語係り受け解析器 Cabocha を利用する。最終文節を述語として抽出した後、最終文節とリンクするすべての文節を取り出し、係助詞「は」、格助詞「が」、「も」を持つ文節を主語として抽出する。上記例文の場合は「学生でした」を述語として、「私は」を主語として抽出される。

英語のセンテンスの主語と述語の抽出には、英文の構文解析器 Stanford Parser を用いる。解析した依存構造の結果から nsubj（名詞句

主語) タグが付与されている語を主語として抽出して、主語から依存関係にある語を述語とし、主語から依存関係にある語まで動詞、助動詞、コピュラを含むフレーズを主語述語フレーズとして抽出する。ここでは“I”は主語として、“I”から“student”の間にある“am a college student”は主語述語フレーズとして抽出される。構文解析器による日本語および英語の主語述語抽出の正確さについては以前の研究で確認している(川村 2018)。

3.2. 日英対照文型と正誤判定基準

抽出した日本語のセンテンスの主語と述語について、これまでに構築した日英対照文型を用いて分類する(川村ほか 2018)。抽出条件に当てはまる主語が複数あるセンテンスに対しては分類を行わない。

表 1 文型と判定基準

文型 1	私 + (は / が / も) + 述語 (動詞) (I)
文型 2	主語 + (ある / いる) (am / is / are / be / have / has / exist / exists)
文型 3	主語 + 名詞 + (です / である / だ) (am / is / are / be)
文型 4	主語 + 形容詞 + (です / φ) (am / is / are / be)
文型 5	主語 + (思う / 考える) (think / believe / consider / guess / suppose / assume)
文型 6	主語 + (知る / わかる) (know / get to know / understand / find / notice / realize / recognize)
文型 7	主語 + (名詞 + できる / 動詞 + ことができる / 動詞 + られる) (can Verb / be able to Verb / am able to Verb / is able to Verb / are able to Verb)

日本語の文型は、中学校・高校の教科書を

参考にして、初級学習者にとって重要な英語文法項目に対応する日本語文型一覧から選定している。日英文型の対照関係から、計 7 種類の文型に対して、主語ベースの誤用判定基準を 1 個、述語ベースの判定基準を 6 種類 24 個考案した。

表 1 には、文型と判定基準一覧を示している。各文型について、上段の日本語は文型で、下段の英文は誤用判定基準である。上段の文型に対して、下段の英文が抽出された英語のセンテンスの主語、または主語述語フレーズになればそのセンテンスに誤りがあると判定している。文型 1 は主語ベースの誤用判定基準であるが、現段階では日本語の主語が「私」である場合に限定している。対応する英語のセンテンスの主語が「I」でなければ誤りと判定する。文型 2 から文型 7 は述語ベースの判定基準であるが、各文型に対して、表 1 に明記した現在形・肯定形のもの以外に、現在形・否定形、過去形・肯定形、過去形・否定形の判定基準も設けている。

日本語の主語が複数検出されたセンテンスに対しては、これらの文型に適応せず、英語のセンテンスに主語があるかどうかのみ判定を行う。また、日本語の主語が 1 つのみ、または無主語のセンテンスで、これらの文型に該当しないものは誤り判定しないこととした。

4. 結果の表示

図 2 は誤り検出結果の一部を表示している画面である。この画面では、入力した全文について対応関係にある日本語作文および英語作文を一行ずつ表示している。英語作文は本システムにより主語部分に誤りがあると判定された場合には主語部分が赤字で表示され、述語部分に誤りがあると判定された場合には述語部分が青字で表示されるように設計されている。ただし、黒字で表示された英語作文であっても正用文とは限らない。それは該当する文型が存在せず、判定できない文も含まれているためである。

また、本システムは誤りの検出結果を表示するとともに、英語作文の詳細情報も確認できる機能を備えている。詳細情報を確認したい英語作文にカーソルを当てると、その英語

作文があるセルが赤色に変わりクリックすることができる。そこでクリックすると画面が英語作文の詳細情報画面に切り替わる。図 3 は詳細情報画面の一部を示している。この画面では、日本語センテンスの主語と述語、英語のセンテンスの主語と述語、システムが判定した文型および英語のセンテンスの主語または述語の誤り判定結果を確認することができる。補足機能として、文型は、便宜上、使

用しているコードがそのまま表示されるため、そのコードにカーソルをあてることで文型情報（品詞カテゴリーおよび文法カテゴリー）が表示されるようになっている。

誤りの検出結果集計画面では、入力した英語作文のうち表 1 に示すそれぞれの文型に当てはまる英語のセンテンスの数、および誤りと判定された英語作文の数の集計結果を表示している。

英作文エラーの検出結果

習文書は主語部分のエラー、熟文書は述語部分のエラーを示しています。
各文の誤り箇所は英文をクリックしてください。

表示の切り替え

集計結果

	日本語	英語
1	なぜなら、レストランは食事をするところだからである。	The restaurant is because it is eating.
2	初のアルバイトはファーストフードショップでした。	The first part-time is fast food shop.
3	私は大学生がアルバイトをすることはとても大切なことだと思います。	I think very important that college students have a part time job.

図 2 誤り検出結果の画面

詳細情報

2	日本語	初のアルバイトはファーストフードショップでした。
	英語	The first part-time is fast food shop.

文法タイプの詳細はカーソルを当ててください。

	日本語	英語	文法タイプ	判定結果
2	主語	アルバイトは	part-time	unknown
	述語	ファーストフードショップでした	is fast food shop	Error

図 3 詳細情報の画面

5. システムの利用

本システムを用いて、実際の日本語エッセイから自由英作文を産出するライティング授業で得られた学習者の作文を分析した。

5.1. 分析用の作文について

学習者は大学生の 1 年生 68 名である。大学生の英語習熟度はヨーロッパ言語共通参照枠 (CEFR) の A1 (Breakthrough) であり、初級レベル学習者に相当する。作文の執筆条件として、日本語作文では「大学生にとってアルバイトをすることは重要である」、または「国内では全てのレストランで喫煙が完全に禁止される

べきだ」のどちらかのテーマを選択し、文字数は 400 字、制限時間は 30-40 分、辞書の使用は不可とした。また、英作文では日本語作文を翻訳し、制限時間は 50-60 分、辞書の使用は不可とした。さらに、学習者は問題に直面すると誤りを犯しそうな構造や語彙を避けるコミュニケーション・ストラテジーをとる (Tarone 1980) 可能性があるため、授業では日本語作文を書く際に後に英語に翻訳することを告知していない。授業で得られた日本語作文・英作文は各 110 エッセイ、各 1499 文であった。

システムを用いて分析するに先立ち、前処理

として英語作文 1499 文から日本語作文と対応していない文、直接話法を含む文、そして命令文は除外して計 1001 文を対象とした。

5.2. 結果

表 2 はシステムによる判定結果を見やすくしたものである。1001 文の学習者の日本語作文は、主語が複数抽出された文(複数主語)が 138 文、抽出されなかった文(無主語)が 366 文、主語が 1 つのみ抽出された文(単一主語)が 497 文となることを明らかにしている。「複数主語」の場合は、英文の主語の有無のみ判定したため、138 文のうち主語がない英文は 9 文であることを「エラー数」欄から読み取れる。

表 2 システムによる判定結果の集計

文の種類	文の数	エラー数
複数主語	138	9
無主語	366	
文型5現在肯定	208	106
文型3現在肯定	17	10
文型5過去肯定	10	9
文型3過去肯定	2	2
文型7現在肯定	1	1
文型7過去肯定	1	1
文型5現在否定	1	1
文型3現在否定	2	1
文型2現在否定	1	1
文型2現在肯定	1	0
単一主語	497	
文型2現在肯定	53	28
文型3現在肯定	59	17
文型5現在肯定	60	13
文型2現在否定	7	7
文型7現在肯定	13	6
文型4現在肯定	10	5
文型5過去肯定	4	4
文型3過去肯定	2	2
文型2過去肯定	1	1
文型3現在否定	1	1
文型6過去肯定	3	1
文型4現在否定	1	1
文型1	97	13

「無主語」と「単一主語」の日本語文の文型とそれぞれの文型に該当する文の数は「文の種類」欄と「文の数」欄に表示している。これら

の文型に対して判定した誤りのある英文の数は「エラー数」欄に「無主語」と「単一主語」について降順に並べ替えて示している。「現在肯定」、「現在否定」、「過去肯定」、「過去否定」は現在形・肯定形、現在形・否定形、過去形・肯定形、過去形・否定形を表している。

表 2 からは、日本語文に対する文型の分類により、今回分析対象とした日本語エッセイには、単一主語の文はおおよそ全体の半分を占め、完全に分類対象外とした複数主語の文は 15%以下であることが分かる。また、無主語の文の中に、表 1 の文型に該当する文は、無主語文の 66%を占めているが、最も使われた文型が文型 5 の現在形・肯定形であり、これに対応した英文エラーが最も多いことが分かる。つまり、「主語 + (思う / 考える)」という日本語の表現を正しく英語に訳すことができないことを意味する。単一主語の中、表 1 の述語ベースの文型に該当した文は比較的少なく、単一主語文の 43%を占めるが、文型 2 の現在形・肯定形における英文のエラーが最も多くなっている。文型 2「主語 + (ある / いる)」に関する英語の表現を習得していないことを示唆している。また、文型 1 は唯一の主語ベースの文型であるが、「私」を使った 97 文のうち、「I」がない文が 13 文あることがわかる。

表 3 エラーと判定されたセンテンスの例

No.	文型	主語	述語	主語英語フレーズ
1	文型5現在肯定		思います	have
2			思います	it's too have
3	文型2現在肯定	レストラン	ある	wall hold
4	文型2現在肯定	ことが	あります	I dislike
5	文型5現在肯定	私は	考える	they smoke
6	文型6過去肯定	私は	知りました	It start

エラーとして検出されたセンテンスの具体例を表 3 に示す。システムの詳細画面で確認できる情報の一部である。ここの主語述語フレーズは前述のように主語から依存関係にある語までの動詞、助動詞、コピュラを含むフレーズである。No.1 と No.2 は無主語文であり、述語にエラーがあると判定されている。No.3 と No.4 は単一主語文であり、述語にエラーがあると判定された文である。No.5 と No.6 は同じ単一主語文であるが、主語、述語ともにエラーがある文

として検出されている。

5.3. 考察

前述の結果を踏まえて、本システムの利用により、教師が詳細な添削を行う前に、学生の基本的な構文についての習得状況を容易に把握できると言えよう。

自由英作文はその量と内容の多様さから、添削は教師にとって大変負担を強いる作業である。本システムによって基本的な構文エラーを自動検出し、赤字で示すことによって、教師による添削時間の短縮が可能となる。また、日本語を文型に分類することによって、学習者の日本語エッセイに使われた文型一覧を確認できるだけでなく、それに対応する自由英作文で使用した英文表現の一覧も確認できる。それぞれの表現について、検出した誤りの集計により、それぞれの表現についての習得状況を把握することができる。学習者別に集計結果を提示することで、添削の個別指導の支援にもなる。

他方、上記の支援を実践的に行うためには、自由英作文の多様さに対応した多くの英語表現をカバーできる日英対照文型および誤りの判定基準の構築は大変重要となる。本システムに取り込んだ文型は、中学校・高校の教科書を参考にして作成したもので、カバーできる英語表現はまだ初級・中級レベルに留まっている。また、結果表示のインターフェイスに関しても、学習者別の表示機能は現在構築中であり、教師によるシステムの評価も今後の課題となる。

6. まとめ

本研究では、教師の英作文添削を支援することに目標とし、学生が日本語エッセイから自由英作文を産出するライティング授業において、日英対照文型に基づいて構築した判定基準を用い、英語の基本的構文となる主語と述語の誤り検出システムを構築した。

そして構築したシステムを利用して実際の学生の自由英作文を分析した。その結果、文型の分類により、基本的な英語表現の中、学習者が作文に使用したものを容易に把握でき、主語と述語に誤りのある英文を検出し、集計することによって、それぞれの表現についての学習者の習得状況を確認できることを明らかにした。

謝辞

本研究の一部は学術研究基金助成金基盤研究(C) (課題番号: 17K01081) の助成を受けている。

参考文献

- Ferris, D. R. (2004). The “grammar correction” debate in L2 writing: Where are we, and where do we go from here?(and what do we do in the meantime...?). *Journal of second language writing*, 13(1), 49-62.
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistics perspective*. Harlow, UK: Pearson Education.
- 川村晃市 (2018). 構文解析器を利用した主語・述語抽出の検討. 神戸大学国際文化学会『国際文化学』, 31, 51-64.
- 木村博是, 木村友保, 氏木道人 (2010). リーディングとライティングの理論と実践. 大修館書店, 東京.
- Koichi Kawamura, Harumi Kashiwagi, Min Kang (2018). An Approach toward Automatic Error Detection in Learners' English Writing Based on the Source Language. *Proceedings of The Tenth International Conference on Mobile, Hybrid, and On-line Learning*, 62-65.
- 太田裕子・佐渡島紗織 (2013). 「自立した書き手」を育成するライティング・センターのチューター研修とチューターの意識: 早稲田大学における実践事例と PAC 分析. 『Waseda Global Forum』, 9, 237-277.
- Tarone, E. (1980). Communication strategies, foreigner talk, and repair in interlanguage. *Language learning*, 30, 417-431.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language learning*, 46(2), 327-369.
- 吉田弘子 (2010). 「大学ライティングセンターに関する考察: その役割と目的」. 『大阪経大論集』, 61, 99-109.