

国立国語研究所学術情報リポジトリ

生成AIはどのように学習者作文の誤用を訂正するのか：誤用分析のアノテーション付与のための一考察

メタデータ	言語: Japanese 出版者: 研究社 公開日: 2026-07-17 キーワード: ChatGPT, プロンプト, 自動作文評価, 訂正フィードバック 作成者: 加藤, 敬介 メールアドレス: 所属: 山梨英和大学非常勤講師
URL	https://doi.org/10.15084/0002000636

This work is licensed under a Creative Commons
Attribution-NonCommercial-ShareAlike 4.0
International License.



第12章 | 生成AIはどのように学習者作文の誤用を訂正するのか

—誤用分析のアノテーション付与のための一考察—

日本語教育の作文の授業において、学習者の書いた作文の添削は日本語教師にとって大きな負担である。本章では、ChatGPTを用いて、学習者作文の誤用訂正を半自動化する4段階のプロセスを提案した。これにより、教師の添削フィードバックの負担が軽減されるだけでなく、誤用訂正のための研究データ整備を安定して行えるようになると思う。

KEYWORD

| ChatGPT | プロンプト | 自動作文評価 | 訂正フィードバック |

□1 生成AIによる誤用訂正と学習者コーパス

日本語学習者作文の添削は、誤用検出からレベル別の修正案提示まで多大な時間と労力を要し、教師の大きな負担となっている。近年、この負担軽減策として大規模言語モデル (LLM) を搭載した生成AIを教師添削の支援・補完に活用する試みが進んでいる。一方で、生成AIは学習者の日本語熟達度を超える修正や、大幅な書き換えにより誤用箇所そのものを消してしまうおそれもある。学習者コーパスとしての再利用や教育的フィードバックの観点からは、生成AIの支援を受けつつも、テキストを適切に区切って「どこを」「どのように直したか」を追跡でき、かつ過剰修正 (直し過ぎ) を抑える指示が不可欠である。

本章ではChatGPTを用い、学習者コーパス収録予定の12編の作文を半自動で処理する4段階のプロセス (以下、本プロセス) を提案する (次ページの表1)。ステップ1・ステップ2に対応する専用ソフトは存在するものの、人手による修正が必要で、現時点で全体の自動連鎖処理は困難である。このため、人手介入を前提に半自動で運用する。

表1 生成AI支援下での誤用訂正のステップ

ステップ	主要処理	ChatGPTの役割
1	長単位分割	BCCWJ長単位境界の自動認定
2	レベル判定	作文から執筆者のJLPT水準を推定する。
3	制約付き訂正	各長単位で誤用検出・日本語レベルに即した修正案を生成する。
4	制約解除リライト	長単位の制約を外し、全文を自然な日本語ヘリライトする。

本プロセスにおいて焦点となる課題は、次の二つである。

- (1) 生成AIによる長単位アノテーションのためのプロンプト設計
- (2) 生成AIの過剰修正（直し過ぎ）を抑制するプロンプト設計

ここで、**プロンプト**は生成AIへの指示文、**アノテーション**は注釈付与を指す。本プロセスでは後述の「長単位」でテキストを区切り、修正前後の比較・分析を可能にする。(1)について、誰が行っても同じ結果になるように安定的に誤用訂正を行い、研究用のデータとして、また、学生向けの教育的フィードバックとして再利用するには、「どこを」「どのように」直したかが追跡可能となるようにテキストを一定の単位で区切り、修正を行う必要がある。本章の分析対象である12編の作文が収録される『日本語学習者縦断作文コーパス (W-CoLeJa)』では、意味の最小単位である形態素の単位で区切った形態素解析データが提供される予定であるが（本書第10章参照）、形態素という単位は誤用訂正には粒度が細かすぎるため、本章では『現代日本語書き言葉均衡コーパス (BCCWJ)』の長単位を採用する。

長単位は文節を内容語と機能語に分割したもので、それぞれを構成要素とするため、まとまった統語的機能を持つ単位として扱いやすいとされる。長単位の境界認定には通常、専用ソフトが用いられるが、本プロセスでは生成AIだけでどこまで正確に実行できるかを検証する。なお、専用ソフト、生成AIのいずれにおいても、学習者コーパスのテキストが誤用・逸脱を含むことによる長単位境界認識の誤りが不可避であるため、人手による修正が必須となる。

(2)について、生成AIはしばしば学習者の語彙選択や語順を大幅に置き換え、誤用箇所そのものを消失させる「直し過ぎ」を引き起こす。日本語教育においては、学習者の執筆した原文を可能な限り尊重し、学習者の日本語熟達度に合わせて最小限の手直しに抑えるのが原則である。そこで本章では過剰修正を抑制する対策として、生成AIに学習者の日本語熟達度を推定させたうえで、レ

ベルに応じた訂正フィードバックを行い、かつ大幅な書き換えを抑制するためにテキストの長単位を保持したままの修正だけを許可するプロンプトを設計し、どのような効果を上げるかを検証する。

□2 先行研究

2.1 長単位と解析ツール

長単位の認定基準は小椋ほか (2011) が詳細に示している。短単位列を長単位へ変換するツールとしては「Comainu」(小澤ほか 2014) が開発され、長単位境界と品詞推定の精度を高めた。近年はspaCyベースの長単位解析系(松田ほか 2022)が登場し、Comainuを上回る精度を示している。

2.2 自動作文評価とレベル判定

一般的に学習者コーパスでは執筆者の熟達度情報が開示されるが、例外もあるため、本プロセスでは生成AIにレベルを判定させる。作文の自動評価については、Li and Liu (2024) が『多言語母語の日本語学習者横断コーパス (I-JAS)』を用いて比較した結果、ChatGPT (GPT-4) が評価精度でBERT、日本語特化LLM (Open-Calm)、従来システム (Jess・jWriter) を上回るが、精度はプロンプト設計に大きく左右されると指摘した。

2.3 生成AIによる訂正フィードバック

英語教育の先行研究では、生成AIが教師の訂正フィードバックをどこまで代替できるかが検討されている。Loem *et al.* (2023) はChatGPT (GPT-3) が教師ありTransformerを上回り、プロンプト次第で修正量を調整できると報告した。Lin and Crosthwaite (2024) はChatGPT (GPT-4) が表層訂正、メタ言語的説明や全文リライトを教師より多用し、過剰・冗長訂正が目立つと指摘した。El Ebyary and Shabara (2024) もChatGPT (GPT-3.5 Turbo) で類似の結果を示し、迅速・網羅的だが過剰訂正が課題と報告している。

□3 生成AI支援下での誤用訂正のステップ

3.1 ChatGPTのモデル選択

本研究ではChatGPTの推論特化モデルGPT-o3 (以下「GPT-o3」) を使用した。ChatGPTには目的別に最適化された複数のモデルが用意されており、GPT-o3は長文脈の保持と複雑な論理推論に秀でている。この特性が、長単位の境界認

定からレベル判定・誤用検出・リライトまでを一貫して処理するのに最適と判断した。なお、一般向けに広く利用されているGPT-4o（高速・汎用モデル）でも同一プロンプトで検証した結果、実用上問題ないことを確認した。

3.2 プロンプト設計の方針

各プロンプトには処理手順、優先順位、出力形式を明記した。また、メタ説明を抑制し、余計な処理負荷が生じないようにした。

3.3 ステップ1 長単位分割

3.3.1 使用プロンプト

ステップ1のプロンプトは以下のとおりである。プロンプトに示した長単位認定ルールは『国語研コーパスポータル』所載のBCCWJの形態論情報¹から作成した。

【ステップ1 プロンプト】

STEP1 作文を長単位に区切る

【ChatGPTの役割】

あなたはBCCWJ（現代日本語書き言葉均衡コーパス）の公式仕様に基づき、文章を「長単位」に分割する専門家です。

【作業指示】

以下の作文に、長単位の区切りを示すため「|」を挿入してください。

出力は、作文に「|」を挿入した結果のみを返してください。

●説明・改行・語順変更・補足コメントは禁止です。

●句読点の後に勝手に改行を入れるなども禁止です。

【作文】²

【適用ルール】

①文節境界の決定

●助詞・助動詞・複合辞の直後に文節境界を置く（これは最優先で必ず実行）

●固有名詞・動植物名・慣用句は、内部で切らない（助詞等があっても固有名詞の中では切らない）

②文節内での長単位認定（優先順位 A→D）

【Step A：付属語の確定（最優先）】

●形態素解析で品詞が「助詞」「助動詞」「複合辞」と判定された語は、それ単独で1長単位にする。

●これら付属語は、直前の語とは必ず切り離す。

●連続する付属語（例：「に|は」）は品詞境界で分ける。

【例】

●会社|に|は|

●世界|で|

●実務慣行|が|ある|。|

【Step B：固定・連結語のまとめ】

●句読点（、。…等）は単独で1長単位。

●複合語（例：公害紛争処理法）、サ変複合語（例：往復運動する）

●並列・同格連結（例：英語＝日本語間）

●語と同機能の記号連続（例：2,000＝m²）は付属語を含まない範囲で1長単位にまとめる。

1| <https://clrd.ninjal.ac.jp/bccwj/morphology.html>（2025.5.31閲覧）

2| ここに解析対象とする学習者の作文のテキストを貼り付ける。

- すでに確定した付属語境界 (Step A) は絶対に越えない。
- 【サ変動詞の一体化 (必須)】
名詞や副詞に「する」「できる」「なさる」「いたす」が直接続く場合は、切り離さずに1長単位にまとめる。
- 【例】
 - 滞在し | た |
 - 往復運動 = し | ている |
 - きちんと = できる |
- 【Step C : 数量表現】
- 単位語の直後や、数値の直前で長単位境界を置く。
- 【例】
 - 平成 | 15年 | 9月 | 15日 | 午後 | 7時 | 33分 |
- 【Step D : 残余トークン (最後の調整)】
- 上記に該当しない自立語列を、文節や意味の自然な範囲で必要に応じて分割する。
- ③競合時の大原則
 - 付属語境界は不変 (Step A > すべてに優先)
 - それ以外の競合時は、「より大きい結合」を優先してまとめる。
- 【出力形式】
- 本文に「|」を挿入した結果だけを返す。
- 改行・語順変更は禁止。
- 【例】
 - その | 会社 | に | は | 世界 | で | 初めて | の | 実務慣行 | が | ある | 。 |
- 【注意】
- 手順・例示・形式を逸脱する出力は禁止。
- ChatGPTが形態素解析に基づいて品詞判断し、指定ルールに従って正確に区切ること。

3.3.2 GPT-o3による長単位分割と人手修正

第11章でも取り上げた12編の学習者による作文を対象に、上記プロンプトを適用してGPT-o3に長単位分割を出力させた。そのうち、大学1年次の中国の大学の調査協力者 (中国語母語話者学習者) が執筆したCCG011-B1-01-01の作文を対象とした長単位境界の出力はテキスト1のとおりであった。「|」が自動付与された長単位境界である。

テキスト1 生成AIの長単位境界の同定結果 (CCG011-B1-01-01)

去年 | の | 夏休み | に | 従妹 | と | 厦門 | に | 遊び | に | 行っ | た | 。 | わたしたち | は | 家 | から | 車 | で | 行っ | た | 。 | 二時間半 | ぐ | らい | か | かっ | た | 。 | 厦門 | に | 三日間 | 滞在 | し | た | 。 | 初日 | は | 有名 | な | 鼓浪嶼 | に | 行っ | た | 。 | その日 | は | 晴れ | て | 雲 | 一つ | ない | 。 | 私たち | は | 柔らかい | 砂浜 | に | 散歩 | し | た | 。 | 海風 | の | 塩辛い | 味 | が | し | た | 。 | とても | 楽し | かつ | た | 。 | そして | 私たち | は | 島 | の | お店 | を | 見 | に | 行っ | た | 。 | 精巧 | で | 美しい | 手作り品 | を | たくさん | 買っ | た | 。 | 最後 | に | 海洋館 | に | 行っ | た | 。 | 可愛い | 海 | の | 生物 | を | たくさん | 見 | た | 。 | 翌日 | は | 方特 | に | 遊び | に | 行っ | た | 。 | 刺激的 | で | 面白 | かつ | た | 。 | 終わっ | た | 時 | に | 盛大 | な | 花火 | を | 見 | た | 。 | 残念 | ながら | 三日目 | に | 台

風 | が | 来 | た | 。 | 私たち | は | どこ | に | も | 行か | なかつ | た | 。 | 四日目
| に | 家 | に | 帰っ | た | 。 | その | 旅行 | は | 一生 | の | 思い出 | です | 。 |

本稿筆者がBCCWJ検索アプリ「中納言」の長単位検索機能を用いて、GPT-o3が生成した境界を検証したところ、CCG011-B1-01-01では、2箇所で長単位境界の誤認が確認された(表2)。いずれも学習者の誤用に起因する長単位境界の誤認ではなかった。これらの境界誤認に対し人手で修正を行い、正しい長単位分割済みテキストを得た。

表2 テキスト1における生成AIの長単位境界の正誤

行	長単位境界 (誤)	長単位境界 (正)	誤りタイプ
4	その日	その 日	境界欠落
5	散歩 し	散歩し	境界過剰

3.3.3 12編全体の誤り概要

表3は、12編の作文すべてに同一プロンプトを適用して行った長単位分割の境界誤認状況を示したものである。

表3 生成AIによる長単位境界の同定の誤り

番号	作文コード	総境界数	境界誤認数	誤分割率 ³⁾ (%)
1	CCG011-B1-01-01	167	2	1.20
2	CCG028-B1-01-01	218	4	1.83
3	CCG034-B1-01-01	243	6	2.47
4	KKB001-B1-01-01	127	11	8.70
5	KKB003-B1-01-01	106	9	8.49
6	KKB011-B1-01-01	143	7	4.90
7	TT007-A1-01-01	203	4	1.97
8	TT010-A1-01-01	189	0	0
9	TT015-A1-01-01	121	3	2.48
10	VVC026-B1-01-01	163	1	0.61
11	VVC078-B1-01-01	198	11	5.56
12	VVC117-B1-01-01	124	7	5.65
	合計	2,002	65	(平均) 3.25

3| 誤分割率 (%) = 境界誤認数 / 総境界数

誤分割率は0% (TT010-A1-01-01) から8.70% (KKB001-B1-01-01) まで散在するが、平均は3.25%であった。誤分割は5類型(表4)にほぼ集約された。中でも誤分割率が8.70%と最も高かったKKB001-B1-01-01では11件中8件がタイプ1「『た』の前の境界欠落」に該当した。

以上から、ステップ1のプロンプトを適用した場合のGPT-o3による長単位境界同定は、特定パターンで誤りが生じやすいことが確認できる。仮に「『た』の前の境界欠落」の回避を促すようにプロンプトを補強し、当該誤認が無くなると仮定した場合、誤分割率は8.70%から2.36%に低下する。誤分割が発生しなかったケースもあり、生成AIによる長単位分割の精度は一定水準に達していると言えるが、ルールの補強とプロンプトの精緻化による誤分割率のさらなる低減が期待される。

表4 長単位境界の誤りタイプ

タイプ	誤りのタイプ	例
1	「た」(過去)の前の境界欠落	(誤) ました→(正) まし た (誤) でした→(正) でし た (誤) った →(正) っ た
2	「ている」「ています」の過剰分割	(誤) て いる →(正) ている (誤) て い ます→(正) てい ます
3	「名詞+する」(サ変動詞)の過剰分割	(誤) 見学 する→(正) 見学する (誤) 見学 し →(正) 見学し
4	数詞+単位の過剰分割	(誤) 三年 前→(正) 三年前 (誤) 8時 半→(正) 8時半 (誤) 1858 米→(正) 1858米
5	名詞句(複合名詞・漢語連鎖)境界欠落	(誤) 高校二年生→(正) 高校 二年生 (誤) 「その日」→(正) その 日

3.4 ステップ2 日本語レベルの判定

3.4.1 判定用プロンプト

誤用修正には学習者の日本語能力を判定することが欠かせない。ステップ2では、日本語能力試験(JLPT)に準拠し、N5・N4・N3・N2・N1以上の5段階でGPT-o3に判定させ、判定根拠を50字以内で提示させた。Zero-shotの簡便なプロンプトで判定させており試行的なものである。実際の運用では自動作文評価研究の知見に基づき評価指標を設計する必要がある。

【ステップ2 プロンプト】

STEP2 作文執筆者の日本語レベル判定

あなたはJLPTに精通した日本語教育の専門家です。学習者の作文を読み、以下の4観点で総合評価し、「レベル」と「根拠」だけを日本語で出力してください。フィードバックや修正文はこの段階では示さないでください。作文中に長単位区分を示す記号「|」が含まれていても無視し、通常の本文として評価してください。

【評価観点】

1. 語彙の幅と適切さ
2. 文法構造の複雑さと正確さ
3. 語用・文体の適切さ（接続詞・時制・敬語など）
4. まとまり・論理構成（段落構成・一貫性）

【出力形式】

レベル：JLPT（N5 / N4 / N3 / N2 / N1以上）

根拠：50字以内で主要な判断理由

【禁止事項】

上記2行以外は出力しない

i+1 フィードバックや訂正案は記載しない

【作文】⁴

3.4.2 判定結果と考察

12編の作文に対するGPT-o3の判定結果を表5に示す。11編がN4、1編（CCG011-B1-01-01）のみがN3であった。N2・N1以上・N5はなかった。

GPT-o3は、短文主体で助詞誤用が頻出する作文を一貫してN4と評価しており、出力された根拠と評価観点との整合性はおおむね良好と判断する。しかし、N2・N1以上・N5レベルは今回出現せず、判定の分布が中位レベルに集中したことから、境界事例に対する判別力や外的妥当性は追加検証が必要である。

また、判定根拠の字数制限を課した結果、詳細な理由づけが省略される傾向がみられた。将来的には、判定根拠の長さを調整し、自動作文評価に関する先行研究の成果を踏まえて、言語的特徴量を定量的に抽出したうえで、人手評価との比較検証を行うなどの対応が必要であろう。

表5 生成AIによる日本語レベル判定結果

番号	作文コード	判定	判定根拠（AI出力）
1	CCG011-B1-01-01	N3	時制・接続の整合があり、語彙も適切で一貫した構成がみられるため
2	CCG028-B1-01-01	N4	基本語彙中心で構文単純、助詞・活用誤り多く一貫性や接続詞使用は限定的
3	CCG034-B1-01-01	N4	基本語彙と丁寧体中心で通じるが不自然表現や語法誤りが散見し複雑構文に乏しい

4| ここに解析対象とする学習者作文のテキストを貼る。

番号	作文コード	判定	判定根拠 (AI出力)
4	KKB001-B1-01-01	N4	基本文型とN4語彙中心、接続詞少なく誤用多数
5	KKB003-B1-01-01	N4	複文と中級語彙あるが、助詞・語形誤用多い
6	KKB011-B1-01-01	N4	基本語彙と簡単文型中心だが、誤用多数でN3未滿
7	TT007-A1-01-01	N4	基本語彙と単文主体で語順誤り多いが、時制接続詞使用
8	TT010-A1-01-01	N4	基本語彙と単文主体、接続詞少数、活用誤り多い
9	TT015-A1-01-01	N4	基本語彙中心、簡単時制と接続詞使用 N4文法「～時」「～たい」等を正確に運用
10	VVC026-B1-01-01	N4	基本語彙とN4文型たり形使用、文法誤用多いが理解可
11	VVC078-B1-01-01	N4	語彙と文法は基本範囲、助詞誤用多く構文単純
12	VVC117-B1-01-01	N4	語彙基本的、たり形・たことがある等、N4文法使用も誤用多い

3.5 ステップ3 制約付き訂正フィードバック

3.5.1 誤用検出・訂正提示用プロンプト

ステップ3では、「分析可能性の確保」を最優先し、(1) 長単位境界を一切変更しない、(2) 執筆者のJLPTレベルを明示したうえで修正レベルを調整するという二つの条件の下でGPT-o3に訂正案を生成させた。これにより、コーパス処理で再利用できるデータ構造を維持しつつ、学習者の認知負荷を大きく上回ることはない語彙・構文を提示することを狙っている。したがって、本ステップのプロンプトは「長単位境界を維持したままの訂正後作文」と「訂正箇所一覧表」を同時出力させる設計とし、長単位保持(境界の再区切り・統合・分割を禁止)、レベル適合(ステップ2の判定結果に応じた語彙・文法のみ使用)、教育的透明性(訂正理由を一覧表で明示)を実現させるべく、以下のとおりとした。

【ステップ3 プロンプト】

STEP3 長単位ベースの訂正フィードバック
 ChatGPTの役割：
 あなたは優秀な日本語教師かつ日本語コーパス編集者です。
 学習者作文に対し、教育的有用性と分析可能性を両立した訂正フィードバックを行います。
 【入力情報】
 ● 学習者レベル：⁵
 ● 作文本文：⁶
 【出力仕様】
 1. 訂正後の作文
 ● 提示された長単位「|」を保持したまま誤用・不自然表現を訂正する。
 ● 学習者が無理なく理解・習得できる範囲で、次のステップとなる表現に修正すること。

5| ここにSTEP2の判定 (N5/N4/N3/N2/N1以上のいずれか) を記載する。

6| ここにSTEP1で出力した長単位分割済みテキストを貼る。

- 文体（普通体／丁寧体）は原文に合わせ、変更しない。
- 長単位の区切り位置は絶対に変更しない。

2. 訂正箇所一覧表

以下の4項目を表形式で提示する。

｜長単位番号｜元の表現｜訂正後の表現｜訂正理由（次のステップとしての意図・言語的説明）｜
【優先順位と制約】

- 訂正は学習者が無理なく理解・習得できる範囲で、次のステップとなる表現を最優先とする。
- 文法・語彙・表記・談話の自然さは上記と競合しない範囲で考慮する。
- 長単位の区切り位置は絶対に変更しない（再区切り・統合・分割禁止）。
- ネイティブらしい表現への最終調整はSTEP4で行うため、本ステップでは行わない。

3.5.2 生成AIによる訂正結果

CCG011-B1-01-01に関するステップ3の訂正出力は、表6およびテキスト2に記載のとおりである。「長単位を保持したまま訂正する」という指示は遵守されており、修正は次の3点に限定された。

- 助詞「も」加筆および「ない」→「なかった」（長単位番号31）
- 格助詞「に」→「を」への置換（長単位番号39）
- 動詞「行く」→「行ける」（可能動詞）への訂正（長単位番号91）

JLPTは公式のシラバスや語彙表を公開していないものの、出力には複雑な構文や高度語彙は含まれず、修正内容はN3相当の範囲内にとどまっていると判断される。他の11編についても長単位を完全保持したまま訂正が行われ、修正は表記・基本文法に関わる範囲にとどまった。

表6 訂正箇所一覧表 (CCG011-B1-01-01)

長単位番号	元の表現	訂正後の表現	訂正理由
31	雲｜一つ｜ない	雲｜一つ｜も｜なかった｜た	「一つも～ない」の否定構文が自然であり、否定の強調表現を学ぶ機会となる。
39	砂浜｜に	砂浜｜を	「～を散歩する」が自然な語法。N3では動詞と助詞の自然な結びつきに注目させる段階である。
91	行か｜なかつ｜た	行け｜なかつ｜た	可能動詞「行ける」の否定形が文脈に合致。「どこにも行けなかった」という意味の正確な表現へ導く。

テキスト2 生成AIによる訂正後の作文 (CCG011-B1-01-01)

去年｜の｜夏休み｜に｜従妹｜と｜厦門｜に｜遊び｜に｜行っ｜た｜。｜わたしたち｜は｜家｜から｜車｜で｜行っ｜た｜。｜二時間半｜ぐらい｜かかっ｜た｜。｜厦門｜に｜三日間｜滞在｜し｜た｜。｜初日｜は｜有名｜な｜鼓浪嶼｜に｜

行っ | た | 。 | その | 日 | は | 晴れ | て | 雲 | 一つ | も | なかつ | た | 。 | 私 | たち | は | 柔 | らか | い | 砂 | 浜 | を | 散 | 歩 | し | た | 。 | 海 | 風 | の | 塩 | 辛い | 味 | が | し | た | 。 | と | て | も | 楽 | し | かつ | た | 。 | そ | し | て | 私 | たち | は | 島 | の | お | 店 | を | 見 | に | 行 | っ | た | 。 | 精 | 巧 | で | 美 | しい | 手 | 作 | り | 品 | を | た | く | さ | ん | 買 | っ | た | 。 | 最 | 後 | に | 海 | 洋 | 館 | に | 行 | っ | た | 。 | 可 | 愛 | い | 海 | の | 生 | 物 | を | た | く | さ | ん | 見 | た | 。 | 翌 | 日 | は | 方 | 特 | に | 遊 | び | に | 行 | っ | た | 。 | 刺 | 激 | 的 | で | 面 | 白 | かつ | た | 。 | 終 | わ | っ | た | 時 | に | 盛 | 大 | な | 花 | 火 | を | 見 | た | 。 | 残 | 念 | な | が | ら | 三 | 日 | 目 | に | 台 | 風 | が | 来 | た | 。 | 私 | たち | は | ど | こ | に | も | 行 | け | な | かつ | た | 。 | 四 | 日 | 目 | に | 家 | に | 帰 | っ | た | 。 | その | 旅 | 行 | は | 一 | 生 | の | 思 | い | 出 | です | 。 |

3.5.3 制約付き訂正フィードバックに関する考察

(1) レベル適合性と過剰修正の回避

GPT-o3は全12編について執筆者のJLPTレベルに見合った修正を提示した。長単位保持を最優先したため、修正内容に過剰修正や過度の高度化は見られず、負荷調整の目的は達成された。

(2) 長単位保持制約と分析可能性

長単位境界を固定する制約を課したことにより、接続詞の不足や語順再構成の必要性といった談話レベルの要改善点は未解決のまま残った。一方で、訂正後のテキストをそのままコーパス解析に再投入できるという利点は確保された。

(3) 訂正箇所一覧表の教育的透明性

「元の表現→訂正後→理由」を一覧化する方式は学習者の誤用に対するメタ認知を促す。ただし、理由文の粒度や正確性にはばらつきがあった。

以上を踏まえると、ステップ3は「習熟度適合性・分析可能性・教育的透明性」という3点をおおむね満たしているものの、談話的自然さの面では改良の余地がある。したがって、ステップ4で長単位制約を解除したリライトを行うことで、これらの弱点を補完できると考えられる。

3.6 制約解除リライト

3.6.1 プロンプト概要

ステップ4では、ステップ3の訂正済み作文から長単位境界を外して自然な日本語に書き直す。プロンプトには(1)日本語母語話者が違和感なく読めること、(2)文体(普通体/丁寧体)を維持すること、(3)学習者レベルを超える難語・高度構文を避けることを盛り込んだ。

【ステップ4 プロンプト】

STEP4 自然な日本語へのリライト

ChatGPTの役割

あなたは優秀な日本語教師かつネイティブ校正者です。学習者による訂正済みの作文を、自然で滑らか、かつ適切な日本語にリライトします。

【入力情報】

- 訂正済み作文：⁷
- 学習者レベル：⁸

【出力仕様】

- 文脈に即した自然な表現・語順・接続詞・文の流れ・言い回しを適切に調整し、日本語母語話者が使う自然な文章に整える。
- 文体（普通体／丁寧体）は原文に従って保持する。
- 元の意味・話者意図・事実関係は変更しない。

【優先順位と制約】

- 自然で滑らかに伝わる日本語表現を最優先とする。
- 語順・助詞・語彙・接続詞・文末表現などを適切に調整し、日本語母語話者が自然に使う表現に整える。
- 学習者レベルに配慮し、過度に難解・硬い表現や文学的表現は避ける。（例：N3学習者に対して、N1レベルの高度な表現は使わない）
- 不必要な直訳調・冗長な表現・語順の不自然さを省き、整理・言い換えで自然化する。
- 学習者が無理なく理解・習得できる範囲で、次のステップとなる自然な表現を意識して最適化する。

3.6.2 リライト後の作文

CCG011-B1-01-01に対するGPT-o3のリライトはテキスト3のとおりである。原文（テキスト1）の事実関係を保ちつつ主語や接続詞を整理して流れを滑らかにした。他の11編でもレベルに配慮した自然なリライトが得られた。

テキスト3 リライト (CCG011-B1-01-01)

去年の夏休みに、従妹と一緒に厦門へ遊びに行った。家から車で行き、二時間半くらいかかった。厦門には三日間滞在した。

初日は、有名な鼓浪嶼に行った。その日は晴れていて、雲一つない青空だった。私たちは柔らかない砂浜を散歩し、海風の塩辛い味を感じた。とても楽しかった。

そのあと、島のお店を見に行き、精巧で美しい手作り品をたくさん買った。最後に海洋館にも行って、可愛い海の生き物をたくさん見た。

翌日は方特というテーマパークに行った。とても刺激的で面白かった。終わったあとには、盛大な花火も見ることができた。

残念ながら、三日目には台風が来てしまい、どこにも出かけられなかった。四日目に家に帰った。

この旅行は、一生の思い出になった。

7| ここにSTEP3で出力された訂正済作文のテキストを貼る。

8| ここにSTEP2の判定（N5／N4／N3／N2／N1以上のいずれか）を記載する。

3.6.3 リライトに関する考察

(1) 流暢性の向上

長単位制約解除により主語や時制の重複解消や接続詞の整理などが可能となり、流暢性・読みやすさが大幅に改善された。

(2) 語彙・文法レベルの維持

原文の語彙・文法レベルが維持され、過度な高度化は行われていない。

(3) 過剰修正（直し過ぎ）のリスク

本プロセスでは顕著ではないが、制約解除により生成AIが表現や語順を大胆に修正、整理すると、学習者独自の語彙選択やスタイルが薄れたり、誤用部分の訂正案が示されなかったりするリスクがある。

(4) 自然な表現への変換

ステップ3と4の比較により、学習者にとっての「誤用の訂正」から「より自然な表現への変換」という2段階の改善過程が可視化される。

□ 4 まとめ

本章では、「長単位分割→レベル判定→制約付き訂正→制約解除リライト」の4段階パイプラインを設計し、GPT-o3で学習者作文を半自動的に処理する手順を提示した。課題(1)「生成AIによる言語単位アノテーションのためのプロンプト設計」については、適用したプロンプトによる12編の作文に対するGPT-o3の平均誤分割率は3.25%で、長単位分割を一定水準の精度で実行できたと評価できる。課題(2)「生成AIの過剰修正（直し過ぎ）を抑制するプロンプト設計」については、長単位境界を保持したままJLPT水準に準拠して訂正させるプロンプトにより、コーパス解析との互換性を確保するとともに、JLPTに準拠した負荷調整によって訂正・リライトの過度な高度化を抑制した。さらに、長単位の制約を外すリライトを併用する二段階修正により、「誤用の可視化」と「自然な表現への変換」を両立させた。

本プロセスには、レベル判定の精度検証、誤用分類タグの定義とプロセスへの組み込み、API連携による一層の自動化といった課題が残る一方で、生成AIツールにより学習者コーパスのデータを簡便に処理できる可能性を示した。今後は、生成AIと人手チェックを組み合わせ、より実用的な誤用分析ワークフローへ発展させたい。

参考文献

- 小椋秀樹・小磯花絵・富士池優美・宮内佐夜香・小西光・原裕(2011)『現代日本語書き言葉均衡コーパス』形態論情報規程集4版(上), 国立国語研究所内部報告書.
- 小澤俊介・内元清貴・伝康晴(2014)「BCCWJに基づく長単位解析ツールComainu」『言語処理学会 第20回年次大会 発表論文集』582-585.
- 伝康晴・小木曾智信・小椋秀樹・山田篤・峯松信明・内元清貴・小磯花絵(2007)「コーパス日本語学のための言語資源—形態素解析用電子化辞書の開発とその応用—」『日本語科学』22: 101-123.
- 松田寛・大村舞・浅原正幸(2022)「UD Japaneseに基づく国語研長単位解析系の構築」『言語処理学会 第28回年次大会 発表論文集』725-730.
- El Ebyary, K., & Shabara, R. (2024). ChatGPT-generated corrective feedback: Does it do What it says on the tin? *Teaching English with Technology*, 24 (3) : 68-89.
- Li, W. & Liu, H. (2024). Applying large language models for automated essay scoring for non-native Japanese. *Humanities & Social Sciences Communications*, 11 (723) : 1-15.
- Lin, S., & Crosthwaite, P. (2024). The grass is not always greener: Teacher vs. GPT-assisted written corrective feedback. *System*, 127 (2024) : 1-19.
- Loem, M., Kaneko, M., Takase, S., & Okazaki, N. (2023). Exploring Effectiveness of GPT-3 in Grammatical Error Correction: A Study on Performance and Controllability in Prompt-Based Methods. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications* (BEA 2023) : 205-219. Association for Computational Linguistics.

(加藤 敬介)