

広島大学教養教育英語科目における 学生ライティングの自動採点ツール作成に向けて

中 川 篤

広島大学外国語教育研究センター

1. はじめに

英語教育において、リーディングやリスニングのような受容系技能に重点を置いた授業では、ライティングの重要性が十分に認識されていない場合が多い。しかし、近年の研究では、ライティング指導が読解力や聴解力の向上に寄与するだけでなく、言語運用能力全般を強化する可能性が示されている。Graham & Perin (2007) は、メモ作成や要約といったライティング活動が読解力の向上に大きく貢献すると指摘している。また、Task-Based Language Teaching (TBLT) を活用した事例では、リスニングやリーディングで得た情報をライティングで統合する活動が理解力や構成力を高めることが示されており、このようなライティング活動が受容系技能を補完し、言語能力全体の発達を促進する可能性を示唆している (Ellis, 2003; Nunan, 2004; Willis & Willis, 2007)。

一方で、実際の教育現場では、ライティング指導を行うことの難しさが大きな課題となっている。ライティング指導は、教師が文法や構成、内容など多岐にわたる要素を個別に評価・指導する必要があるため、特に大規模クラスでは教師の負担が非常に大きい (Brown, 2007; Ellis, 2009; Goldstein, 2005; Hyland & Hyland, 2006)。広島大学の教養教育英語科目においては、特に受容系技能を主に担当するコミュニケーション IB・コミュニケーション IIB はクラスサイズが大きく、前期コミュニケーション IB の1クラス平均学生数は46.8名に達する (2024年度の筆者の場合)。加えて、教育主担当教員は一週間に10クラスを担当するため、すべての学生に質の高い個別フィードバックを提供することは現実的に著しく困難である (Warschauer & Ware, 2006)。

こうした課題を解決するため、AWE (Automated Writing Evaluation: 自動作文評価システム、以下 AWE) の活用が注目されている。AWE は、近年発展の著しい人工知能 (AI) や自然言語処理 (NLP) の技術を活用し、学生の作文に対して即時的かつ一貫性のあるフィードバックを提供することを目的とするツールである (Shermis & Burstein, 2013; Warschauer & Ware, 2006)。これにより、教師の負担を軽減しつつ、学生が迅速に自身の課題を認識し改善に取り組む環境を提供できる可能性がある。

本稿では、既存の AWE の利点と課題を概観するとともに、要約評価に特化した新たな AWE の実装可能性とその教育的効果について議論する。筆者が開発したプロトタイプシステムは、要約の内容の網羅性、正確性、構造、明確さ、言い換え能力の5つの観点から多角的に評価を行い、学生に個別化されたフィードバックを提供する仕組みを備えている。今後は、このシステムを実際の教育現場で実証実験し、その有効性を検証していく予定である。

2. 先行研究

2.1. ライティング指導の課題

ライティング指導は、英語教育の中でも特に時間と労力を要する分野であり、教育現場では多

くの課題が指摘されている。Hyland & Hyland (2006) や Nation (2009) は、ライティング指導が単に文法や語彙の正確性を評価するだけでなく、文章の構成や一貫性、論理展開といった高次スキルにも対応する必要がある点を強調している。また、Ferris (2003) は、学生一人一人のスキルや目標に応じた個別化されたフィードバックが、指導の質を大きく左右する重要な要素であることを強調している。さらに、Hyland (2003) は、フィードバックが学習プロセスを促進するためには、適切なタイミングで具体的な指摘を行うことが不可欠であると述べている。

しかし、大規模クラスにおいて、効果的なライティング指導を実現することは非常に困難である。この問題はフィードバックの拡張性 (scalability) の欠如に起因しているといえる。Ellis (2009) は、従来の指導方法ではクラス規模が拡大するにつれて教師の負担が指数的に増大し、指導の効率性が著しく低下すると指摘している。特に、フィードバックを必要とする学習者が多数いる場合、この問題はさらに深刻化し、全ての学生に対して適切な個別対応を行うことは現実的に不可能である。本学でも、筆者が担当する受容系科目では1クラスあたりの学生数が40～50名に達することが一般的であり、このような環境では、教師が全員に個別かつ具体的なフィードバックを提供するための時間を確保することは極めて難しい。その結果、指導のタイムリーさを維持することはほぼ不可能となり、学習者にとって効果的なフィードバックが提供されない状況が生じている (Bitchener, 2008)。この問題は、フィードバックの質だけでなく、教師の負担軽減と教育効率の向上を求められる現代の教育現場において、早急に解決が必要な課題である。

さらに、教師が提供するフィードバックの一貫性、すなわち評価者内信頼性 (intra-rater reliability) の問題も課題として挙げられる。同じ教師が作文を評価する場合、その基準が一貫していることが求められるが、Hyland & Hyland (2006) は、短期間で大量の作文を評価する際に評価基準が揺らぐ可能性を指摘している。例えば、評価対象が抽象的な内容を含む場合や、教師が疲労している状況では、評価が主観的に偏るリスクが高まる (Brown, 2007)。また、Warschauer & Ware (2006) も、複数クラスを担当する教師が限られた時間の中でライティング指導を行う際、その作業が非常に負担の大きいものであることを指摘している。その結果、フィードバックが抽象的または一般的な内容にとどまり、学生にとって実用性が低下する場合がある。

このように、大規模クラスでは、フィードバックの量や質だけでなく、タイミング、拡張性、そして評価者内信頼性といった複数の課題が顕在化している。これらの課題を解決するためには、従来の方法に依存するのではなく、指導の効率化と質の向上を両立するための新たなアプローチが必要である。

2.2. 作文評価システム (AWE) の利点

こうしたライティング指導の課題を解決する手段として、近年注目を集めているのが自動作文評価システム (AWE) である。AWE は、人工知能 (AI) や自然言語処理 (NLP) の技術を活用し、学生の作文に対して即時的かつ一貫性のあるフィードバックを提供するシステムであり、特に大規模クラスのような環境で、従来の教師中心の評価方法の限界を補うものとして期待されている (Shermis & Burstein, 2013; Warschauer & Ware, 2006)。特に以下の3つの利点が指摘されている。

まず、AWE の最大の利点は即時性である。学生が作文を提出した直後にフィードバックを受け取ることが可能であり、これにより学習者はタイムリーに課題を認識し、改善に取り組むことができる (Hyland, 2003)。特に、大規模クラスでは、教師が全員に対して迅速なフィードバックを提供することは難しいため、AWE の即時性は教育効率を大幅に向上させると考えられる。

次に、AWE は一貫性と客観性を提供する。アルゴリズムに基づいて評価を行うため、評価者の主観に左右されず、一貫した基準で作文を評価することができる (Shermis & Burstein, 2013)。これは、評価者内信頼性の課題を軽減するだけでなく、公平性の確保という点でも重要である。

さらに、AWE は単なるエラー検出にとどまらず、教育的支援としての機能も備えている。たとえば、文法や語彙のアドバイスだけでなく、文章の可読性や論理展開の改善提案を行うことが可能であり、学生が自律的に学習を進めることを支援する (Attali & Burstein, 2006)。これにより、教師はフィードバック作業の負担を軽減し、指導の検討・改善などのより創造的な業務に時間を割くことができる。

代表的な AWE として、Grammarly や Criterion (ETS) などが挙げられる。Grammarly は主に文法やスペリングのチェック、スタイルの改善に焦点を当てている一方、Criterion は教育機関向けに開発され、作文の構造や内容に重点を置いた評価を行う。このように、AWE はライティング指導の効率化と質の向上に大きく貢献するツールである。

2.3. 既存の作文評価システム (AWE) の課題

AWE には、大規模クラスにおけるライティング指導の課題を解決する可能性を持つ技術である。Attali & Burstein (2006) によれば、AWE は文法や語彙の正確性を即時かつ一貫して評価できる点で有用性が高い。しかしながら、その限界も指摘されている。

例えば、Hyland & Hyland (2006) は、AWE が文章の一貫性 (coherence)、結束性 (cohesion)、論理的展開といった高次スキルの評価において限界があると述べている。さらに、Granger (2009) は、AWE が比喩や皮肉、多義語といった文脈依存的な要素を正確に解釈できないことを指摘している。

また、学習者個々のニーズに応じた柔軟なフィードバックが提供できない点も課題である。Ferris (2003) は、一般化されたフィードバックが学習者にとって具体性を欠く場合、学習効果を十分に引き出せない可能性があるとして述べている。このような課題は、AWE が教育現場で完全に適応するために克服すべき重要な障壁である。

2.4. 先行研究のまとめ

先行研究から、ライティング指導は英語教育の中でも特に時間と労力を要する分野であり、大規模クラスにおいてはフィードバックの提供にさまざまな課題が生じることが明らかになった。ライティング指導においては、文法や語彙の正確性だけでなく、学生一人一人に応じた個別化されたフィードバックの提供が重要である (Ferris, 2003; Hyland, 2003; Hyland & Hyland, 2006; Nation, 2009)。一方で、クラス規模が拡大するにつれて教師の負担が指数的に増加し、指導の効率性が著しく低下し、学習者にとって効果的な指導が提供されない状況が生じている (Bitchener, 2008; Ellis, 2009)。このように、大規模クラスでは教師によるフィードバックの量的・質的な限界が顕著である。

このような状況下で活躍する可能性があるのが自動作文評価システム (AWE) である。本研究は、これらの課題を克服し、リーディングやリスニングのような受容系技能に重点をおいた大規模クラスの授業におけるライティング指導の効率と効果を向上させる AWE を作成することを目的としている。具体的には、大規模クラスにおける要約スキルの育成を対象とし、以下の目標を達成することを目指す：

- (1) 拡張性の向上：大規模クラスでも即時的かつ一貫性のあるフィードバックを全ての学生に提供し、教師の負担を軽減する。
- (2) フィードバックの一貫性と客観性の確保：自動化されたアルゴリズムによる評価を用いることで、フィードバックの質のばらつきを最小限に抑える。
- (3) 高次スキルの評価：言い換え能力や要約の内容的正確性、文章構造の適切性などを多角的に評価し、具体的で建設的なフィードバックを提供する。

要約を対象とした理由は、文章の要点整理や理解の深化を図れることに加え、受容系技能に重点を置く授業で筆者が要約活動を実施しているためである。

3. システム開発

3.1. システムの設計原理

本システムは、学生の要約スキルを効率的かつ多角的に評価し、適切なフィードバックを提供するために設計されている。特に、要約タスクにおいて重要な情報選択の網羅性と内容の正確性、そして言い換え能力を重視した評価基準を採用している。また、教育現場での大規模クラス運営に対応し、教師の負担軽減と学生の学習効果向上を目的としている。本システムは3.3で示す4つのモジュールで構成されており、それぞれに必要な入力ファイルと評価項目を関連付けて動作する。

3.2. 評価手法とフィードバック生成

内容評価においては、自然言語処理技術を活用し、以下の手順で分析を行う：

- (1) 前処理：形態素解析による文章の正規化
- (2) 特徴抽出：TF-IDF ベクトル化による重要語の特定
- (3) 類似度計算：コサイン類似度による内容の網羅性評価
- (4) 構造分析：文の長さや論理展開の評価

フィードバックは以下の3段階で生成される：

- (1) 評価スコアの算出：各評価基準に基づくスコアリング
- (2) 履歴分析：過去の学習データとの比較
- (3) フィードバック選択：評価結果と学習段階に応じた適応的フィードバック

3.3. システムの概要

3.3.1. 要約分析モジュール (SummaryAnalyzer)

要約分析モジュールは、学生が作成した要約を多角的に評価するために設計されている。このモジュールの目的は、要約の質を客観的かつ定量的に測定し、学生に対して有益なフィードバックを提供することにある。特に、要約タスクにおいて重要な「内容の網羅性」「内容の正確性」「構造」「明確さ」「言い換え能力」という5つの観点から、要約文のパフォーマンスを分析する。

内容の網羅性は、学生の要約が原文の重要な情報をどれだけカバーしているかを評価する項目である。この評価では、教師が事前に作成した要点ファイル (summary_key_points.txt) を基準に、学生要約がこれらの要点を含んでいるかを測定する。要約と要点リストを TF-IDF ベクトル化し、

コサイン類似度を用いて内容の網羅度を数値化することで、学生が情報選択能力を適切に発揮しているかを判断する。以下は内容の網羅性を評価する核となるアルゴリズムである：

```
def analyze_content_coverage (self, summary: str, original_text: str) -> float:
    # TF-IDF vectorization を用いて文書類似度を計算
    tfidf_matrix = self.vectorizer.transform ([original_text, summary])
    cosine_sim = cosine_similarity (tfidf_matrix[0:1], tfidf_matrix[1:2]) [0][0]
    return cosine_sim
```

内容の正確性は、要約が原文の意図を正しく反映しているかを評価する。単に要点を含むだけでなく、それらが正確かつ誤解なく表現されているかを重点的に測定する。また、文法エラーやスペルミスが正確性に与える影響も考慮し、TextBlob ライブラリを使用して文章の品質を補完的に評価する。

構造の評価は、要約文の論理的な一貫性や段落構成の適切性を測定するものである。平均文長を基準とし、文の長さが適切な範囲（15語前後）から大きく逸脱していないかをチェックすることで、文章構成品力を測定する。また、段落内の論理的な流れや結束性についても評価し、分かりやすい文章の構築ができているかを分析する。

次に、明確さは、文章が文法的に正確で、読み手にとって明瞭であるかを評価する項目である。文法チェックツールを用いて句読点の誤用や大文字の不適切な使用などの機械的エラーを検出し、これらの問題が明確さに与える影響をスコア化する。

最後に、言い換え能力は、学生が原文をそのまま抜粋せず、適切に要約を行う能力を測定する項目である。この評価では、N-gram 分析や文レベルの類似度計算を活用し、原文と要約文の過剰一致を防ぎながら、要約文のオリジナリティを評価する。この指標は、剽窃を防止し、学生の独創性を促進するために重要となる。このモジュールは、原文テキスト (original_text.txt) と要点ファイル (summary_key_points.txt) を基準に、それぞれの評価項目の観点から学生の要約の質を定量的に測定し、学生が要約タスクで身に付けるべきスキルを客観的に評価する、本プログラムの中核的なコンポーネントである。

3.3.2. 学習者追跡モジュール (StudentPerformanceTracker)

学習者追跡モジュール (StudentPerformanceTracker) は、学生一人一人の要約スキルの成長を記録し、学習者ごとの進捗や課題を把握するために設計されたモジュールである。このモジュールは、SQLite データベースを利用してデータを管理し、学生の過去の成績を基にした個別指導を可能にする。また、クラス全体のパフォーマンスを可視化し、教師が教育戦略を改善するためのデータ基盤として機能する。

このモジュールでは、学生番号 (student_id)、各評価項目 (内容の網羅性、内容の正確性、構造、明確さ、言い換え能力)、および総合スコアがデータベースに記録される。教師はこれらの情報をもとに、特定の学生の進捗を追跡するだけでなく、クラス全体の成績分布や平均スコアを分析することができる。また、履歴データを時系列で取得する機能を備えており、各学生の過去5回分の評価結果を参照することで、パフォーマンスの改善傾向や停滞を特定し、それらに応じたフィードバックを提供できる。この機能により、学生がどのスキルを向上させる必要があるのか

を明確にし、個別の指導計画を立てる際の参考情報を提供する。

さらに、このモジュールはクラス全体の統計データを生成する機能も持っている。これにより、教師は学生全体の学習状況を把握し、教育の効果を測定することが可能となる。具体的には、各評価項目の平均値計算を通じて、指導の成果や課題を可視化する。このように、StudentPerformanceTracker は、個別学習支援と全体的な教育効果の向上を目指した重要なモジュールである。

3.3.3. 適応型フィードバック生成モジュール (AdaptiveFeedbackGenerator)

適応型フィードバック生成モジュール (AdaptiveFeedbackGenerator) は、学生の評価結果をもとに個別化されたフィードバックを生成し、学生が要約スキルを効果的に改善できるよう支援するために設計されたモジュールである。このモジュールは、教師が設定した評価基準やテンプレートを活用し、各学生に対して具体的かつ実用的なアドバイスを提供する。

このモジュールでは、まず、教師が作成した要約サンプル (sample_summaries.txt) や要点ファイル (summary_key_points.txt) をもとに、学生にとって理想的な要約の方向性を示す。評価結果に基づいて、具体的な改善点や優れた点を明示し、どのスキルに重点を置いて取り組むべきかを指摘する。また、過去のパフォーマンスデータを参照し、学生ごとの成長や改善点を分析することで、フィードバックに個別性と適応性を持たせている。

さらに、このモジュールは、評価項目ごとに異なるテンプレートメッセージを生成する仕組みを備えている。例えば、「内容の網羅性」については、特定の keypoints が不足している場合に、それを補う具体的なアドバイスを提供する。一方、「言い換え能力」については、原文との過度な一致が見られる場合に、言い換えの重要性を強調するフィードバックを提供する。このように、AdaptiveFeedbackGenerator は、学生が自身の課題を的確に把握し、次の学習ステップに進むための支援を行う。

このモジュールの特徴的な点は、履歴ベースの分析を活用していることである。フィードバックの生成は、現在の評価結果とパフォーマンス履歴の両方を考慮して行われる。各評価基準について、スコアを「高」「中」「低」の3段階に分類し、それぞれに対応する5つの具体的なフィードバックテンプレートを用意している。さらに、過去の評価履歴との比較分析により、改善が見られた項目や更なる注意が必要な項目を特定し、学習者の動機付けを高めるポジティブなフィードバックと建設的な改善提案を組み合わせ提供している。これにより、AdaptiveFeedbackGenerator は、教師の負担を軽減しながら学生にとって効果的なフィードバックを提供し、学習プロセスを継続的に支援する役割を果たしている。

3.3.4. ユーザーインターフェース (SummaryFeedbackGUI)

ユーザーインターフェース (SummaryFeedbackGUI) は、本システムの操作性を高めるために設計された直感的な GUI (グラフィカルユーザーインターフェース) である。このインターフェースは、教師が効率的に評価プロセスを管理し、学生にフィードバックを提供できるようサポートする。

GUI は複数のタブで構成されており、それぞれが特定の機能に対応している。メインタブでは、学生の要約を読み込み、評価を実行し、フィードバックを生成する基本的な操作を行うことができる。設定タブでは、原文テキスト (original_text.txt) や要点ファイル (summary_key_points.txt)、要約サンプル (sample_summaries.txt) などの必要なファイルを指定し、評価に用いるパラメータ

を調整することが可能である。また、統計タブでは、クラス全体の成績をグラフ形式で表示し、学生の進捗状況を視覚的に把握できる機能が提供されている。

さらに、GUIには重み調整機能も備わっており、教師は評価基準の重み付けを動的に変更できる。この機能により、教育目的や指導方針に応じて柔軟にシステムをカスタマイズすることが可能である。また、フィードバックのプレビュー機能を活用することで、生成されたフィードバックを事前に確認し、必要に応じて修正を加えることができる。

このユーザーインターフェースは、教師が効率的に複数クラスや大量の要約を処理するための直感的で使いやすい設計となっており、本システム全体の運用性を向上させている。特に、データの視覚化機能やリアルタイムのフィードバック生成機能により、教師が迅速に教育的意思決定を行えるよう支援している。また、本システムは、マルチスレッディングを活用して大量の要約を効率的に処理することができる。これにより、大規模クラスでの運用にも対応可能となっている。エラー処理やログ記録も実装されており、システムの安定性と信頼性を確保している。

本システムの完全なソースコードと実装の詳細は、以下で公開している：

<https://drive.google.com/drive/folders/1RY5LyfsDMHX4EYO6bh4k1pG25M84kAfb?usp=sharing>

4. 今後の課題と展望

本研究で開発した AWE の有効性を検証し、さらなる改善を図るため、以下の取り組みを計画している。

4.1. 実証実験による効果検証

実際の教育現場での実証実験を通じて、システムの教育的効果を検証する。具体的には、本システムの教育的効果を検証するため、準実験デザインによる実証研究を計画している。対象は広島大学教養教育英語科目の前期コミュニケーション IB を受講する学生 ($N = 90$) とし、システムを使用する実験群と従来の教師フィードバックを受ける統制群を設定する。評価指標として、要約スキルの向上度（事前・事後テスト）、学習動機の変容（質問紙調査）、およびシステムの有用性（半構造化インタビュー）を測定する。これにより、本システムが学習者の要約スキル向上に与える影響を定量的・定性的に分析する。

4.2. システムの改善と機能拡張

現行システムの限界を踏まえ、以下の改善を計画している：

(1) 談話構造の解析機能

- (ア) 結束性 (cohesion) と一貫性 (coherence) の自動評価
- (イ) 修辞構造理論 (RST) に基づく文章構造分析

(2) 言い換え評価の高度化

- (ア) 意味的類似度に基づく言い換え表現の質的評価
- (イ) 文脈を考慮した適切性判定

これらの改善により、より包括的な要約評価が可能となる。

4.3. 教師支援機能の充実

教育現場での実用性を高めるため、以下の機能追加を予定している：

- (1) クラス全体の学習傾向分析ダッシュボード
- (2) 個別学習者の成長過程の可視化ツール
- (3) カリキュラム設計支援機能

4.4. 技術的課題への対応

システムの安定性と精度向上のため、以下の技術的改善を行う：

- (1) 評価基準の自動調整機能の実装
- (2) 大規模言語モデルの活用による評価精度の向上
- (3) 処理速度の最適化と拡張性の改善

4.5. 汎用性の検証

本研究で開発したフレームワークの他分野への応用可能性を検討する。具体的には、アカデミックライティングやプレゼンテーション原稿の評価への適用を想定している。これにより、英語教育における包括的な自動評価システムの構築を目指す。

5. 終わりに

本研究では、広島大学教養教育英語科目における大規模クラスのライティング指導の課題を背景に、自動作文評価システム（AWE）を開発し、その設計原理、システム構造、教育的有効性について検討した。本システムは、学生の要約スキルを効率的かつ多角的に評価し、即時的で一貫性のあるフィードバックを提供することで、クラスサイズが大きくなりがちな受容系技能を扱う授業においてもライティング指導を実施可能にする可能性を持つ。

今後は実証実験を通じて、本システムが学生の学習成果や動機付けに与える影響を検証し、より実用的で柔軟性のあるツールへと改良を進めていく予定である。また、他分野や他技能への応用を検討することで、教育現場全体の効率化と学習者支援の充実を図りたいと考えている。

最終的には、本システムを通じて、教師と学生の双方にとって負担が少なく、効果的な教育環境を実現することを目指す。教育現場からのフィードバックを積極的に取り入れ、技術的改良と教育効果の向上を追求し続けることで、ライティング教育の新たな可能性を切り開いていきたい。

参考文献

- Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® v.2.0. *The Journal of Technology, Learning, and Assessment*, 4(3), 1–30.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102–118. <https://doi.org/10.1016/j.jslw.2007.11.004>
- Ellis, R. (2009). Corrective Feedback and Teacher Development. *L2 Journal*, 1(1), 3–18.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Routledge.
- Granger, S. (2009). The contribution of learner corpora to the study of second language acquisition. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 13–32). John Benjamins. <https://doi.org/10.1075/scl.33.02gra>

- Hyland, K. (2003). *Second language writing*. Cambridge University Press.
- Hyland, K., & Hyland, F. (2006). *Feedback in second language writing: Contexts and issues*. Cambridge University Press.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. Routledge.
- Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation: Current applications and new directions*. Routledge.
- Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research*, 10(2), 157–180. <https://doi.org/10.1191/1362168806lr190oa>
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.

ABSTRACT

Toward the Creation of an Automated Scoring Tool for Student Writing in Liberal Arts English Courses at Hiroshima University

Atsushi NAKAGAWA

Institute for Foreign Language Research and Education

Hiroshima University

Recent studies have shown that writing instruction not only improves reading and listening skills but also enhances overall language proficiency. However, in large classes, providing high-quality, individualized feedback to all students poses a significant challenge due to time constraints. To address the issue, this study developed an Automated Writing Evaluation (AWE) system tailored for summary-writing tasks. Using natural language processing technologies, the system evaluates summaries across five dimensions: content comprehensiveness, accuracy, structure, clarity, and paraphrasing skills. Key features include content evaluation via TF-IDF vectorization and cosine similarity, paraphrasing assessment using N-gram analysis, and an adaptive feedback mechanism informed by learners' performance history. Future research will involve empirical testing with students to evaluate the system's educational effectiveness. This study aims to facilitate effective writing instruction in large classes while improving teacher efficiency.