

# 学習者作文評価システム「jWriter」による習熟度と論理性の自動評価

李在鎬<sup>1,5</sup>, 長谷部陽一郎<sup>2</sup>, 伊集院郁子<sup>3</sup>, 青木優子<sup>4</sup>, 村田裕美子<sup>5</sup>

<sup>1</sup>早稲田大学 <sup>2</sup>同志社大学 <sup>3</sup>東京外国語大学 <sup>4</sup>東京福祉大学 <sup>5</sup>ミュンヘン大学

jhlee.n@gmail.com, yhasebe@mail.doshisha.ac.jp, ijuin@tufs.ac.jp, yuko172@gmail.com, y.murata@lrz.uni-muenchen.de

## 概要

本研究グループでは、日本語学習者の作文執筆を支援する目的で、「jWriter」

(<https://jreadability.net/jwriter/>, [8][10]) というウェブアプリケーションを開発・運用している。2016年から日本語学習者が書いた作文の習熟度を自動評価するシステムとして公開してきたが、2022年秋には、論理性を自動評価する機能を追加した。論理性の自動評価は、[11]に基づく回帰式に加え、評価項目の一部を診断的評価として提供しており、文章の内容に応じて推敲のためのヒントを提示する機能も備えてある。

## 1 はじめに

私たちは科研費による補助のもと、2016年度より日本語学習者の文章執筆を支援するウェブシステム「jWriter」を開発している(図1)。



図1 jWriterの初期画面

図1のテキストボックスに文章を貼り付け、「実行」ボタンをクリックするだけで、自動評価を行うと同時に、文章が持つ様々な情報を可視化する機能を持っている。このシステムに関わるプロジェクトの概要は、以下のとおりである。

### 【第1期】

- いつ：2016年4月～2019年3月
- 何を：科研費 基盤研究(C)  
(<https://kaken.nii.ac.jp/ja/grant/KAKE NHI-PROJECT-16K02794/>)の補助を得て文章の能力値を自動評価する「jWriter」を開発・公開した。
- 誰が：李在鎬(研究代表者, データ分析), 長谷部陽一郎(研究分担者, システム構築), 村田裕美子(研究協力者, データ収集)
- なぜ：日本語学習者の作文執筆を支援するため。
- どうやって：日本語学習者コーパス「多言語母語の日本語学習者横断コーパス(I-JAS)」[6]に含まれている自由作文のデータを学習データとして利用し、線形回帰分析による習熟度の計算モデル[11]を作成し、ウェブアプリケーションで公開した。
- どこで：  
<https://jreadability.net/jwriter/>

### 【第2期】

- いつ：2019年4月～2023年3月
- 何を：科研費 基盤研究(B)  
(<https://kaken.nii.ac.jp/ja/grant/KAKE NHI-PROJECT-19H01273/>)の補助を得て文章の論理性を自動評価する機能を「jWriter」に追加した。
- 誰が：李在鎬(研究代表者, データ分析), 長谷部陽一郎(研究分担者, システム構築), 伊集院郁子(研究分担者, データ作成), 青木優子(研究分担者, データ作成), 村田裕

美子（研究協力者，データ収集）

- なぜ：日本語学習者の作文執筆を支援するため．特に文章を論理的に書くことの重要性を注意喚起するため．
- どうやって：「現代日本語書き言葉均衡コーパス（BCCWJ）」  
(<https://clrd.ninjal.ac.jp/bccwj/>)に収録されている教科書データと J-STAGE(<https://www.jstage.jst.go.jp/>)で公開されている学会要旨を学習データとして利用し，線形回帰分析による習熟度の計算モデル[10]を作成し，ウェブアプリケーションとして公開した．
- どこで：  
<https://jreadability.net/jwriter/>

第1期目は，文章の習熟度（入門，初級，中級，上級，超級）を自動評価するシステムを作ることに注力した[5][10][12]．

第2期年からは，文章の論理性を自動評価する機能を追加することを目指している．2017年から行っている私たちの試みは，プレイメントテスト[7]において活用する例もあるし，伊集院郁子氏が代表を務める科研費プロジェクト[2]に連携し，アカデミック・ライティングの教育支援ツールとしての活用を計画している例もある．今後，日本語教育の学習・教育インフラの一つとして活用されることを目指している．

## 2 習熟度の判定

習熟度とは，言語教育における「Proficiency」[3]に相当するもので，いわゆる初級，中級，上級といった段階性に基づいて記述される概念である．伝統的にプロフィシエンシーは話す力と関連づけられることが多いが，本プロジェクトでは，日本語学習者の「書く力」に注目した．高等教育機関における学習活動において，「書く力」は不可欠だからである．こうした背景のもとで2016年に日本語学習者コーパス「多言語母語の日本語学習者横断コーパス（I-JAS）」[6]の作文データを学習データとして利用し，線形回帰分析で5段階の習熟度（超級・上級・中級・初級・入門）を推定する計算モデルを作成した[10][12]．

### 習熟度判定用計算モデル [12]

$$y=1.637+\text{平均文長}\times 0.045+\text{中級後半語}\times 0.021+\text{タイプトークン比}\times -0.430+\text{動詞}\times 0.015+\text{中級前半語}\times 0.011+\text{総文字数}\times -0.004+\text{和語}\times 0.007+\text{漢語}\times 0.007 \quad (R^2=0.760)$$

## 3 論理性の判定

文章が持つ論理性の度を計算モデルで推定し，可視化する機能を「jWriter」に実装した．具体的には，文章の論理性のかなめである接続詞を抽出し，カテゴリー別に表示する機能を実装したり，計算モデルに基づいて論理性を4段階（優・高・中・低）で自動評価する仕組みを入れたりしている[11]．

論理性の自動評価のために教科書コーパスと論文要旨コーパスを多変量解析の方法で分析し，計算モデルを作成している．重回帰分析の精度を示す決定係数では78%の精度で論理性の度合いを判定できるようになっている．

論理性判定のための学習データとして「現代日本語書き言葉均衡コーパス（BCCWJ）」に含まれる教科書データと J-STAGE で公開されている学会要旨を使用した．これらのデータを「優」「高」「中」「低」の4レベルに分け，計量テキスト分析の方法で，論理性を計算するモデルを作成した．なお，学習データを4レベルに分けたのは，[9]の分析結果を参考にしたからである．

### 論理性判定用計算モデル[11]

$$y=3.099+(\text{普通名詞率}\times -2.837)+(\text{初級前半語率}\times 4.292)+(\text{接続詞率}\times -49.773)+(\text{動詞率}\times 4.281)+(\text{外来語率}\times -6.17)+(\text{代名詞率}\times -10.525)+(\text{助動詞率}\times 0.279)+(\text{形状詞率}\times -0.691) \quad (R^2=.780)$$

## 4 実装

ここでは，「jWriter」の実装について述べる．「jWriter」は Ruby によるウェブアプリケーションである Sinatra を用いて構築されている．基本的なサーバーサイド処理の流れは次のとおりである．

1. 習熟度判定に必要な最低入力文字数を満たしているか確認

2. 入力されたテキストを文に分割
3. 各文に UniDic を用いた形態素解析を施して語彙情報を取得
4. 同時に「日本語教育語彙表」に基づいて各語彙項目の難易度情報を取得
5. 習熟度判定計算モデルを用いて習熟度を判定
6. 論理性判定に必要な最低入力文字数を満たしているか確認
7. 論理性判定用計算モデルを用いて論理性判定
8. 習熟度と論理性についての診断的評価を出力

1 と 6 の最低入力文字数については、下回ると正確な判定が難しいことがわかっており、習熟度判定については 300 字以上あることが条件となっている。したがって「jWriter」では 300 字未満のテキストを評価することはできない。論理性判定については本来であれば 2,000 字以上のテキストであることが望ましいが、学習者の作文としてそこまでの長さを求めることは難しい場合が多いため 1,000 字以上を条件としている。

2 の文分割については、意味的な分析ではなく純粹に形式的な条件に基づいて行っている。まずテキストを段落に分割するが、その際には「改行+全角スペース」「改行+半角スペース」「2 個以上の改行文字の連続 (= 行空け)」のいずれかを改段落のマーカ―として用いている。学習者の作文テキストは必ずしも定型的でなく、文の途中などで改行されていることが珍しくない。不注意などでそうなっている場合もあるが、段落をあらためる意図なしに行の右端で改行を入れる形で執筆を行う学習者も多いとみられる。そこで、意図的な改階段とそうでない改行を区別するために上記の条件でテキストを段落に分割している。その後、各段落を文に分割するが、その際には「。」「.」「?」「!」をマーカ―として用いる。鉤括弧類が使用されているときなど、実際にはこれらのマーカ―を伴わない文も存在するが、括弧類の使い方は入力されるテキストによって様々な相違があるため（閉じ括弧の直前で句点

を挿入する／挿入しないなど）「jWriter」では文分割の際に括弧の情報は用いていない。

3 の形態素解析については、UniDic-2.1.2 [13] を形態素解析器 MeCab を通じて利用している。

UniDic を利用する理由としては主に次のことが挙げられる。

- ・短単位での形態素解析は、誤りを含んだ表現や非慣例的な組み合わせの表現が含まれやすい学習者テキストの分析に適している
- ・基本形、活用形、品詞に関する情報に加えて「和語」「漢語」「外来語」など語種に関する情報を取得できる
- ・学習者による日本語テキストを分析するのに十分な情報量と性能を有している

4 の処理で用いる日本語教育語彙表[14]は過去の科研費グループ「汎用的日本語学習辞書開発データベース構築とその基盤形成のための研究」で開発された資源で、日本語教育で用いられる語彙項目を 6 段階のスケールに分類している

(<http://jhlee.sakura.ne.jp/JEV/>)。UniDic から得られる語彙情報と合わせて日本語教育語彙表の語彙レベル情報を用いることで、学習者による日本語テキストに含まれるほとんどの語について必要な情報を得ることができる。

ここまでの処理を経て、5 で実際の習熟度判定が行われる。判定結果は回帰式の値に応じて「入門」「初級」「中級」「上級」「超級」の 5 段階に分類される。回帰式の値がどのレベルの値域にも含まれない場合には「測定不可（入門以下のため）」もしくは「測定不可（超級以上のため）」が出力される。また入力テキストが 1,000 字を満たしていればさらに論理性判定が実行される。論理性判定のレベルは「優」「高」「中」「低」の 4 段階であり、「優」を超える値については「測定不可」と判定される。

「jWriter」では習熟度の判定が行われると、結果のレベルに応じて「語の使い方に関する診断的評価」として「語の多様性」「漢語力」「長文作成力」「難解語」に関するアドバイスが示される。この機能は、回帰式を開発するために実施したデータ分析で用いた同じレベルグループのテキストが持つ特徴の平均値と、入力されたテキストの特徴値を比較することで実現している。ユーザーの作文能力の具体的な課題を特定する機能であり、画面上には各項目の値域ごとに用意されたアドバイスのテキストと共に、グループ平均と実際の値とを示したグラフが表示される。これらのフィードバックを活用することで、学習者や教師は現在の作文能力に関する詳しい情報だけでなく、今後の学習・指導につながる気づきを得ることができる。

## 5 「jWriter」が目指す評価

「jWriter」は自動採点システム[1]である。つまり、評価のツールということになるが、本システムはどのような考え方のもとで作られた評価ツールであるかについて述べる。

伝統的に評価研究においては、妥当性 (validity) と信頼性 (reliability) という指標が重要視されている[4]。これらは評価ツールの良し悪しを捉える上で必ず議論になる観点である。妥当性は、評価しようとする能力を正しく評価できているのかに関連する指標である。信頼性は、評価結果の一貫性・安定性の指標である。

「jWriter」が重視したのは、評価の信頼性である。というのも、「jWriter」は誰が、どこから、いつ評価しても、同じ文章であれば、同じ結果を出力するからである。人間であれば、同じ文章であっても、評価者がかわれば、評価もかわるのが一般的で、評価の一貫性や安定性を担保するために多くのコストがかかる[1]。しかし、「jWriter」のような自動採点においては、評価の一貫性や安定性は（計算モデルが同じであれば）不変である。

評価の妥当性には、結論的というと現時点では判断が難しいと考えている。というのは、そもそもコンピュータは回答の求め方が人間とは異なるからである。前節で述べた習熟度の評価のための公式に注目してほしい。平均文長や語種の頻度をもとに、

「jWriter」は習熟度を評価しているが、人間はこうした指標を使って文章を評価するとは考えにくい。

人間の評価では、様々な文構造を使い、自然なつながりにおいて自身の考えを述べているか、語彙の使い方が適切か、さらに主張をとまなう文章であれば根拠がきちんとしているかなどの視点が用いられる。「jWriter」は人間が行っているこれらの評価指標を間接的に捉えることを目指している。つまり、文構造の適切さや文と文のつながりの良さを平均文長や動詞や接続詞の使用率で捉え、語彙使用の適切さは和語や漢語の使用率で捉えることを目指しているのである。このように「jWriter」が用いている指標はいずれも間接的なものであるため、その妥当性に関しても現時点において宣言的なことは言えない。今後、「jWriter」を用いた実践研究が進むことにより、明らかになることであると考えている。

このように妥当性に関しては十分に明らかでない現状においても「jWriter」は有用な評価ツールであると考えている。それは、教育現場に求められる様々な評価活動を考えた場合、評価の妥当性より信頼性やレスポンスの速さが優先される場面があるからである。典型的にはプレースメントテストのような場面や学習者が自身の文章を推敲する場面である。これらの場面では、即時に反応してくれる仕組みが必要だからである。妥当性に関して、多少の問題があったとしても教師が介入することで修正が可能であると考えている。このことは、教師の介入が前提にされない評価活動においては「jWriter」を使うのは望ましくないことを意味する。例えば、入学試験のような個人の人生を決めるようなハイ・ステークスな評価において、現状のシステムでは利用が難しい。

## 謝辞

本研究は JSPS 科研費 19H01273, 19K21637, 22H00667 の助成を受けたものである。

## 参考文献

- [1] 石井雄隆・近藤悠介(編). 英語教育における自動採点-現状と課題, ひつじ書房, 2020.
- [2] 伊集院郁子. 日本語母語話者と日本語学習者の作文に見られる接続表現の比較, 2022 年度日本語教育学会秋季大会予稿集, pp. 65-70.

- [3] 鎌田修・嶋田和子・迫田久美子. プロフィシエンシーを育てる～真の日本語能力をめざして, 凡人社, 2008.
- [4] 近藤ブラウン妃美. 日本語教師のための評価入門, くろしお出版, 2012.
- [5] 小森和子・李在鎬・長谷部陽一郎・鈴木泰山・伊集院郁子・柳澤絵美. 教師による評価とコンピュータによる自動評価はどの程度一致するのか, 2018 年度日本語教育学会秋季大会予稿集, pp. 278-283.
- [6] 迫田久美子・石川慎一郎・李在鎬 (編). 日本語学習者コーパス I-JAS 入門 : 研究・教育にどう使うか, くろしお出版, 2020.
- [7] 柳澤絵美・小森和子・李在鎬・長谷部陽一郎. 日本語オンライン・プレイスメント・テストの開発, 2018 年度日本語教育学会春季大会予稿集, pp. 304-309.
- [8] 李在鎬・長谷部陽一郎・村田裕美子. 学習者作文の習熟度に関する自動判定と Web システムの開発について, ICT×日本語教育, pp. 38-53, ひつじ書房, 2019.
- [9] 李在鎬. BCCWJ の学校教科書コーパスの計量的分析 —日本語教育のためのリーダビリティと語彙レベルの分布を中心に—, 計量国語学 32(3), pp. 147-162, 2019.
- [10] 李在鎬. 書くことを支援する自動評価システム「jWriter」(特集 AI や ICT が変える言語教育), 日本語学 2021 冬号 40(4), pp. 42-51, 2021.
- [11] 李在鎬・伊集院郁子・青木優子・長谷部陽一郎・村田裕美子. 論理的文章の自動評価に関する研究—アカデミック・ライティングへの貢献を目指して—, 2021 年度日本語教育学会春季大会予稿集, pp. 223-226.
- [12] Lee, J., and Hasebe, Y.: Quantitative Analysis of JFL Learners' Writing Abilities and the Development of a Computational System to Estimate Writing Proficiency, *Learner Corpus Studies in Asia and the World* 5, pp. 105-120, 2020.
- [13] 伝康晴・小木曾智信・峯松信明・小椋秀樹・内元清貴・山田篤・小磯花絵. 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, pp. 101-123, 2007.
- [14] Sunakawa, Y., Lee, J. and Takahara, M.: The Construction of a Database to Support the Compilation of Japanese Learners Dictionaries, *Acta Linguistica Asiatica* 2(2), pp. 97-115, 2012.