

言語指標を用いた日本語文章の難易度の可視化と教育への応用

劉 婧怡¹ 内田 諭²

¹九州大学大学院地球社会統合科学府 ²九州大学大学院言語文化研究院

liujingyi0416@gmail.com uchida@flc.kyushu-u.ac.jp

概要

本研究は、日本語学習者向けの文章の難易度判定の可視化と日本語能力試験（JLPT）との対応付けを行うものである。別の研究で提案した分類精度が 88.0% の 35 の言語指標集合を基に、難易度を文字語彙、文構造、談話構造の 3 観点から捉え、それぞれリッジ回帰モデルを構築し、スコアを可視化した。JLPT の読解文章を用いて提案手法を検証した結果、提案モデルは全体的に判定精度が高く、観点別の難易度スコアを科学的根拠に基づいて提示できる点で、大規模言語モデルを用いた判定手法に対して優位性を示した。さらに、日本語教育における難易度と教育枠組みの関連性を明らかにするため、一般の日本語文章と JLPT の読解文章における観点別難易度の相対位置を可視化した。これらの成果は学習者のレベルに合わせた読解教材の作成・開発に貢献することが期待される。

1 はじめに

文章の難易度は、読みやすさや理解のしやすさを評価するための重要な指標である。難易度に関する研究成果は外国語教育の分野において、学習者のレベルに適した教材の選定や作成に活用されている。この研究領域は、英語を対象とした研究が起源とされているが、近年では日本語学習者向けの研究も見られるようになった ([1], [2], [3], [4])。これらの先行研究では、日本語の学習資料を基に、文章全体の難易度を判定する手法が提案されてきたが、本研究は特に以下の 2 点に着目する。

A. 文章難易度判定の多角化

文章の難易度は、使用されている語彙、文構造、結束性・一貫性など複合的な要素によって決定されるため、文章全体の難易度は必ずしもこれらすべての要素が同一の重みであるとは限らない。例えば、初級レベルの文章であっても高度な文構造が使用される可能性があり、中級レベルの文章においても上

級レベルの文構造が多用される場合がある。本研究では、劉[4]が提案した日本語文章難易度判定に最適な言語特徴集合を基盤とし、文字語彙、文構造、談話構造の各側面に着目することで、文章の多角的な難易度判定結果の提示を試みる（研究課題①）。

B. 教育枠組みとの関連性の明確化

難易度の判定基準は、研究者の独自の定義に従い分類基準（初級、中級、上級等）が設けられることが多く、言語運用能力を評価する統一的指標との整合性確保が課題となっている。教育の文脈において最大限の有用性を発揮するためには、難易度レベルと一般的な教育的枠組みの関係性を明確化することが重要である。そのため本研究では、受験者数が世界的に最も多い日本語能力試験のレベルを参考レベルとし、文章のレベル評価において、「初級」や「中級」などの判定結果に加え、日本語能力試験のレベルを付与することにより、レベルイメージの具体性を向上させる（研究課題②）。

2 関連研究

学習者向けの日本語文章の難易度判定の課題を解決するために、劉・内田[1]は、旧日本語能力試験の 1 級から 4 級の 4 つのレベルを判定基準として、語彙と構文の難易度の言語指標を用いて難易度推定手法を提案し、分類精度は 72.2% であったと報告している。しかし、使用された基準データは 10 年以上前のものであり、現在の日本語教育の現状に適していない可能性がある。川村・北村[2]は、語彙の難易度と構文の複雑さに着目し、教科書、専門書、白書を基に、文章の難易度を日本語能力試験の 5 段階レベルに N0 レベルを追加した 6 段階に分類し、テストデータにおいて 79% の精度を達成した。また、李[3]は、日本語教科書と『現代日本語書き言葉均衡コーパス』を用い、日本語教育関係者や専門家の判定によってテキストを初級前半から上級後半の 6 段階に分け、平均文長、漢語率、和語率、動詞率、助詞率の 5 つの指標を用いて重回帰分析を行

った。これらの研究では線形回帰の手法を使用しており結果は解釈しやすいが、取り上げられる言語指標が語彙と文構造関連のものに限定されている。一方、劉[4]は、日本語学習用の教科書文章に基づき、5つの言語単位（文字、語彙、文節、文、談話）の86の言語特徴を抽出し、判定モデルの構築および性能の評価を行った。また、言語教育分野で求められるモデルの簡潔性および説明性を確保するために、特徴選択の手法で文章の難易度と最も関連性の高い35の要因を取り上げ、88%の予測精度を達成した。しかし、この研究では難易度が3クラス（初級～上級）に分類されているため、言語教育現場で有用なより詳細な難易度分類が課題として残されている。

近年の日本語文章の難易度研究では、大規模な基準データと詳細な言語指標を使用することで、文章全体の難易度判定における分類精度が向上している。しかし、ユーザーフレンドリーな形での情報提供や統一された難易度基準が設けられておらず、実際の日本語教育現場での活用が難しい可能性がある。本研究では、劉[4]が提案した学習者向けの日本語文章難易度の判定モデルを基盤とし、選定された最適な言語特徴集合を基に、文章の難易度を文字語彙、文構造、談話構造の各側面から多角的に評価する。さらに、文章の各側面の難易度を日本語能力試験のレベル別読解文と関連付けて可視化することで、難易度レベルを日本語学習者や言語教師にとって親しみやすい教育枠組みと結び付ける。

3 観点別の難易度判定モデルの構築

3.1 基準コーパス

本研究では、独自に構築した『日本語教科書コーパス』を基準とする。このコーパスは、2010年以降に日本で出版された日本語学習用教科書から構成され、教科書に示された難易度に基づいて初級・中級・上級の3レベルが明確に区分されている。対象教科書は計40冊（初級21冊、中級11冊、上級8冊）で、収録テキスト数は1,252編（初級753編、中級325編、上級174編）、総語数は約43万語である。このコーパスにより、近年の日本語教育の実態をより正確に反映できることが期待される。

3.2 言語指標の選定

言語指標は、劉[4]が選定された学習者向けの日本語文章の難易度判定の結果を基に、35の言語指

標からなる最適な言語特徴集合を採用した。さらに、それらを文字語彙、文構造、談話構造の3つの側面を分類した。指標と定義は付録Aで示す。

3.3 構築手法と手順

本研究では観点別の難易度判定モデルを以下の手順で構築する。まず、教科書文章のレベルと観点別の35の言語指標の値を抽出し、標準化を行う。次に、標準化された言語指標を説明変数、教科書文章の難易度を従属変数として、3つのリッジ回帰モデルを作成し、各言語指標の重み係数を求める。観点別の難易度スコアは、個々のモデルにおいて重み係数と標準化された言語指標の積を合算することで算出される。

リッジ回帰は、線形回帰の一種であり、損失関数に偏回帰係数の2乗和（L2正則化項）を用いてモデルの複雑さを制御し、汎化能力の向上を図る手法である。リッジ回帰モデルを採用した理由は、説明変数間の多重共線性への対処である[5]。本研究では、文構造面の最大依存距離、依存距離標準偏差、平均文長、平均文節数の4つの言語指標においてVIF（分散拡大要因）が10以上であったため、リッジ回帰モデルを使用した。これらの指標は難易度判定に重要であると考えられ、35の言語指標はForward stepwise selectionの手法により自動的に選定された。

3.4 結果と考察

表1は、観点別のリッジ回帰モデルの重み係数を示している。重み係数は実験によって得られた結果であり、必ずしも実際の言語的直感と一致するとは限らない。リッジ回帰は、特徴間の相関関係を考慮しながら、全体的な予測誤差を最小化するように係数を推定する。そのため、ある特徴の係数が負になっても、それは他の特徴との関係性の中で解釈する必要がある。

判定モデルの評価指標として、平均二乗誤差（MSE）と調整済み決定係数（調整済み R^2 ）を使用する。MSEはモデルの予測精度を評価し、値が小さいほど精度が高いことを示す。調整済み R^2 は、説明変数の数を考慮してモデルの説明力を評価し、値が大きいほど説明力が高いことを示す。表1によると、MSEは文字語彙から談話構造まで数値が上昇する傾向が読み取れ、その中でも文字語彙面の予測性能が最もよく、談話構造に向かって徐々に低下

することを示している。また、調整済み R^2 の数値から、文字語彙、文構造、談話構造の言語特徴は、それぞれ約 6 割、5 割、3 割未満の文章難易度を説明できる。2 つの評価指標の値は十分に高くないことから、文章の難易度は 1 つの観点の言語特徴だけでは説明できないことが示唆される。

表 1 観点別の判定モデルおよび評価

文字語彙	重み係数	文構造	重み係数
Guiraud値	0.233	最大依存距離	0.187
中級後半語率	0.141	依存距離標準偏差	0.007
助詞率	0.141	平均依存距離	0.102
漢字率	0.142	平均文長	0.542
固有語率	-0.058	平均文節数	-0.425
中級前半語率	0.038	文節数	0.075
動詞_複合動詞率	0.069	名詞節含有率	0.052
漢語率	-0.07	談話構造	
動詞率	-0.013	隣接する文間平均類似度	0.209
助詞_格助詞率	-0.156	対比・選択型接続詞率	0.073
延語数	0.040	Topic3 (確率)	0.017
助詞_接続助詞率	0.095	Topic1 (確率)	0.012
受身含有率	-0.014	逆接型接続詞率	-0.073
和語率	-0.058	補足・説明型接続詞率	0.126
上級前半語率	0.086	Topic5 (確率)	0.001
2級語彙含有率	0.053	隣接する文間類似度標準偏差	0.013
内容語含有率	-0.102		
GKR	-0.025	評価指標	文字語彙
異語率	0.055	MSE	0.223
代名詞率	0.001	調整済み R^2	0.596
			0.475
			0.276

4 モデルの妥当性検証

4.1 検証手法

本研究では、提案した観点別の難易度判定モデルと新規文章データの計算手法の妥当性を 2 つの手法で検証する。検証データには、旧日本語能力試験と新日本語能力試験の読解文を使用する。1 つ目の手法では、3 つのモデルによる観点別難易度スコアが、旧試験データのレベル間の差異が再現可能かを、ウェルチの一元配置分散分析と多重比較で検証する。2 つ目の手法は、『旧試験・新試験認定の目安・認定基準』[6] (以下:『認定基準』) を基に、提案した観点別の難易度スコア算出手法の妥当性を検証する。『認定基準』によると、新試験の N1, N2 は旧試験の 1 級, 2 級に相当し、幅広い場面での日本語理解力を評価する。N4, N5 は旧試験の 3 級, 4 級に対応し、基本的な日本語の理解力を測定する。一方、新試験における N3 レベルは他のレベルと異なる特徴を有し、独立したレベルとして位置づけられる。

各観点の難易度スコアを特徴量として k-means 法によるクラスタリングを行うことで、N3 レベルと他のレベルとの位置関係、および新旧試験間のレベルの対応関係を視覚的に把握することが可能となる。これにより、本手法がレベル間の差異を適切に捉えていることを裏付けられることが期待される。

4.2 検証用データ

検証 1 では 2008 年以前の旧日本語能力試験の読解文を収録した『旧日本語能力試験読解文コーパス』を、検証 2 では、検証 1 のデータを加えて 2012 年と 2018 年の新試験の読解文を収録した『新日本語能力試験読解文コーパス』を使用する。両コーパスとも OCR で電子化し、誤字脱字の確認・修正を行っている。旧試験コーパスには 1 級 137 編, 2 級 145 編, 3 級 48 編, 4 級 65 編の文章が含まれ、新試験コーパスには N1 が 17 編, N2 が 18 編, N3 が 12 編, N4 が 9 編, N5 が 8 編の文章が収録されている。

4.3 検証結果

検証 1 ではウェルチの一元配置分散分析を実施した結果、文字語 (F(3,391)=92.412, $p < 0.01$)、文構造 (F(3,391) = 261.221, $p < 0.01$)、談話構造 (F(3,391)=30.392, $p < 0.01$) のいずれの面においても、レベル間の差が確認された。多重比較の結果、文字語彙面では 1 級と 2 級の間に有意な差が確認されなかったが、文構造面と談話構造面においては隣接する全てのレベル間で有意差が認められた。このことから、本研究で提案した観点別の難易度算出手法は、旧日本語能力試験の各レベル間の差異を有意に検出することができたと言える。

検証 2 では、機械学習ライブラリの scikit-learn(Ver.1.5.2)の cluster モジュールを用いて、エルボー法 ([7], [8]) による最適なクラスター数を 3 と特定し、k-means 法によるクラスターの分割状況およびレベルの分布状況を図示した (付録 B)。文字語彙、文構造、談話構造の 3 つの次元があるが、図の可読性を確保するために各 2 次元で日本語能力試験のレベル間の類似性を考察した。「文字語彙面」と「文構造面」(付録 B) では、レベルの分布が「4 級→N5→3 級 (N4) →N3→2 級→N2→1 級→N1」の順で配置され、日本語能力試験の認定目安に基本的に従っている。レベル間の類似状況を見ると、4 級と N5, 3 級と N4, 2 級と N2, 1 級と N1 がそれぞれ近くに分布している。一方、N3 は全てのレベル

から相対的に独立し、クラスター2の中心に配置しており、『認定基準』に一致している。一方、他の2次元の観点からもN3は相対的に独立していることから、本研究で提案された観点別の難易度スコアの計算方法は、『認定基準』の説明に根拠を提供することができ、観点別の難易度の目安を提供するのに十分な説明力を持つと考えられる。

5 LLMでの判定結果と比較

本節では大規模言語モデル(LLM)の判定結果と比較するため、本研究で用いたデータをChatGPT(APIを使用, プロンプトは付録Dを参照, model=gpt-4o-2024-11-20, temperature=0)を用いてレベル判定の精度を実験した結果(one shot learning), 教科書コーパスでは正答率は約81%となったが, 特に上級での正解率が低くなった(初級88%, 中級77%, 上級57%)。それに対し, 本研究で使用した特徴集合では全体の判定精度が88%であり[4], 初級, 中級, 上級の判定精度はそれぞれ95%, 83%, 65%であったが, いずれもLLMより高いことが明らかになった。また, 日本語能力試験のデータで実験した結果(データ数が限られるためzero shot learningを実施), 正答率は約52%となり, 一つ上のレベルと誤推定してしまう場合が目立った(約46%)。さらに, 観点別の評価(10点満点)ではN4とN5の平均スコアが逆転する, スコアのばらつきが小さく各級の弁別性が低いという結果になった。このように, LLMは一定の判定精度を示す一方, 上級レベルでの判定精度や観点別評価における課題が明らかになった。提案手法はより高い判定精度と観点別の難易度スコアを科学的根拠に基づいて提示できる点で優位性があると言える。

6 JLPTのレベルとの対応付け

前述の通り, 難易度判定基準は教育の枠組みと関連を付けることで, 文章のレベル分けの利点を読解資料の作成やカリキュラムの設定に活かすことができる。本研究では新規文章を用いた難易度数値の可視化において, 文字語彙, 文構造, 談話構造の観点において文章のレベルが新日本語能力試験とどのよう

にマッピングできるかを示す。判定結果をわかりやすく提示するために, レーダーチャートを用いて視覚化する(図1)。新規文章の各観点の難易度スコアを【実線】で, 隣接する日本語能力試験のレベルを【点線】で表示する。試験の各観点の難易度スコアの95%信頼区間を【エリアプロット】で示す。



図1 新規文章の可視化結果

ここでは『できる日本語中級』[9]第10課のテキスト(原文は付録Cを参照)を用いて説明する。この文章は3つの観点から算出された難易度スコアは6.87, 6.88, 6.34ⁱであった。これらのスコアを日本語能力試験のレベル別読解文と比較した結果, 文字語彙面ではN1, 文構造ではN2, 談話構造ではN3~N4に相当することが明らかになった。3観点の平均スコアは6.70であり, 日本語能力試験のレベル別の参考値ⁱⁱと比較すると, 当該文章は総合的にN2からN1レベルの文章であると判定できる。

7 おわりに

本研究では, 学習者向けの日本語文章の難易度判定において, 複数の観点から評価を行う難易度判定モデルを提案し, 一般の日本語文章と日本語能力試験のレベルとの対応関係を可視化して明らかにした。提案手法はLLMでの結果より高い精度と客観的な言語指標に基づく判定を実現した。この研究により, 日本語学習者や教育者は学習者のレベルに適した教材選択, 読解指導, 学習計画の立案, および適切な指導を行うことが可能となる。今後の課題として, 可視化した結果の教室での具体的な活用方法の検討や日本語学習者に対するユーザー調査などが挙げられる。

ⁱ 新規文章の難易度スコアは, 教科書文章の難易度を基準として新規文章の位置を示す。3観点の回帰式により算出された数値は, 相対的に比較するために, 同一の数値範囲に正規化する必要がある。本研究では0から10までの数値範囲を設定する。

ⁱⁱ 参考値の算出方法は新規文章と同じである。N1は7.00, N2は6.57, N3は5.14, N4は4.23, N5は2.63である。

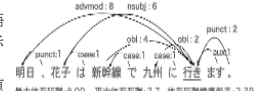
謝辞

本研究は JSPS 科研費 JP23K21949 の助成を受けたものです。

参考文献

- [1] 劉志宇, 内田理. 日本語を学習する外国人を対象とした日本語テキスト難易度推定手法. 研究報告自然言語処理, (11), pp. 1-5, 2012.
- [2] 川村よし子, 北村達也. 日本語学習者のための文章の難易度判定システムの構築と運用実験. Journal CAJLE, 14, pp. 18-30, 2013.
- [3] 李在鎬. 日本語教育のための文章難易度に関する研究. 早稲田日本語教育学, 21, pp. 1-16, 2016.
- [4] 劉婧怡. 学習者向けの日本語文章の難易度判定モデルの構築と最適化: 難易度関連要因の探索および予測効果の解明. 計量国語学, 33(6), pp. 405-420, 2024.
- [5] 竹内秀一. 数値実験による線形回帰における多重共線性の影響力評価. 東京経済大学人文自然科学論集, 132, pp. 3-26, 2012.
- [6] 旧試験・新試験認定の目安・認定基準 (オンライン) (2024年6月1日閲覧)
<https://www.jlpt.jp/about/pdf/comparison01.pdf>
- [7] 安尾萌, 河野雪乃, 山西良典, 松下光範. 化粧動画の工程検索を指向した化粧語彙セット構築の試み. 知能と情報 (日本知能情報フレンジイ学会誌), 35(2), pp. 645-654, 2023.
- [8] Cui, M. Introduction to the k-means clustering algorithm based on the elbow method. Accounting, Auditing and Finance, 1(1), pp. 5-8, 2020.
- [9] 嶋田和子 (監修). できる日本語中級. アルク. 2013.

付録 A 言語特徴の定義・計算手法

観点	特微量	定義・計算手法
文字語彙面	延語数	単語の度数の総和
	Guiraud値	異語数を延語数の平方根で割った値
	異語率	文章中に含まれる異語数を延語数で割った値
	漢字率	漢字数を全文字数で割った値
	中級前半語含有率	『日本語教育語彙表』を参照し、中級前半、後半、上級前半語のそれぞれを延語数で割った値
	中級後半語含有率	
	上級前半語含有率	
	2級語彙含有率	国際交流基金(1994)の『日本語能力試験出題基準』を参照し、2級語彙の数を延語数で割った値
	動詞・複合動詞率	国立国語研究所 (2015)『複合動詞レキシコン』が含まれる語の数を延語数で割った値
	受身含有率	品詞が助動詞の「れる」「られる」「される」の語数を延語数で割った値
	漢語率/固有語率/和語率	漢語数/固有語数/和語数を延語数で割った値
	GRR	外来語数を漢語数で割った値
	内容語含有率	名詞数・動詞数・形容詞数・副詞数の総和を延語数で割った値
動詞率/助詞・格助詞率/助詞・接続助詞率/代名詞率/助詞率	出力される品詞解析の結果による動詞の数/格助詞の数/接続助詞の数/代名詞の数/助詞の数を文章の延語数で割った値	
文構造面	最大依存距離	1文あたりの単語間の依存関係を示す距離の最大値
	依存距離標準偏差	標準偏差・平均値
	平均依存距離	延語数を文数で割った値
	平均文節数	文節数を文数で割った値
	文節数	出力される解析結果による文節として認識された数
	名詞節含有率	出力される解析結果による(名詞節)として認識された数を文節数で割った値 例: 彼女が作った料理を食べた。
	最大依存距離	
談話構造面	隣接する文間平均類似度	BERTモデルによる算出された文章中の隣接する文の間における類似度(余弦類似度)の平均・標準偏差 (BERTモデルはbert-base-multilingual-casedを使用した。)
	隣接する文間類似度標準偏差	
	逆接型(接続詞)	(逆接、補足・説明、対比・選択)を表す接続詞のそれぞれの使用回数を延語数で割った値(接続詞の分類はヒューマン・アカデミー(2017)『日本語教育能力検定試験完全攻略ガイド:日本語教育能力検定試験学習書』第5版を参照)
	補足・説明型(接続詞)	逆接:「しかし」「だが」等 補足・説明:「というは」等
	対比・選択型(接続詞)	
	Topic1(確率)	LDA(Latent Dirichlet Allocation)トピックモデルを使用し、文章のトピックを確率で集計した。各文章におけるTopic1-5の確率(トピック数は、GridSearchCVを使用してperplexityを最小化するように選定された。)
	Topic3(確率)	
Topic5(確率)		

注: 以上の解析結果は、Python3による自然言語処理ライブラリのspaCy(Ver.3.4.4)とGiNZA(Ver.1.2)を用いて行った。自動的に出力されるもので、人間の直感と異なる場合がある。以下の資料を参考にした。
 1. 国立国語研究所. 複合動詞レキシコン (オンライン) (2022年6月閲覧。) <https://vlexicon.ninjal.ac.jp/>
 2. 日本語教育語彙表 (オンライン) (2022年7月閲覧。) <http://jlee.sakura.ne.jp/JEV/>
 3. 国際交流基金, 日本国際教育支援協会 (編). 日本語能力試験出題基準. 凡人社. 1994.
 4. ヒューマンアカデミー. 日本語教育能力検定試験完全攻略ガイド: 日本語教育能力検定試験学習書 (第5版) 翔泳社. 2017.

付録 B 検証 2 の結果

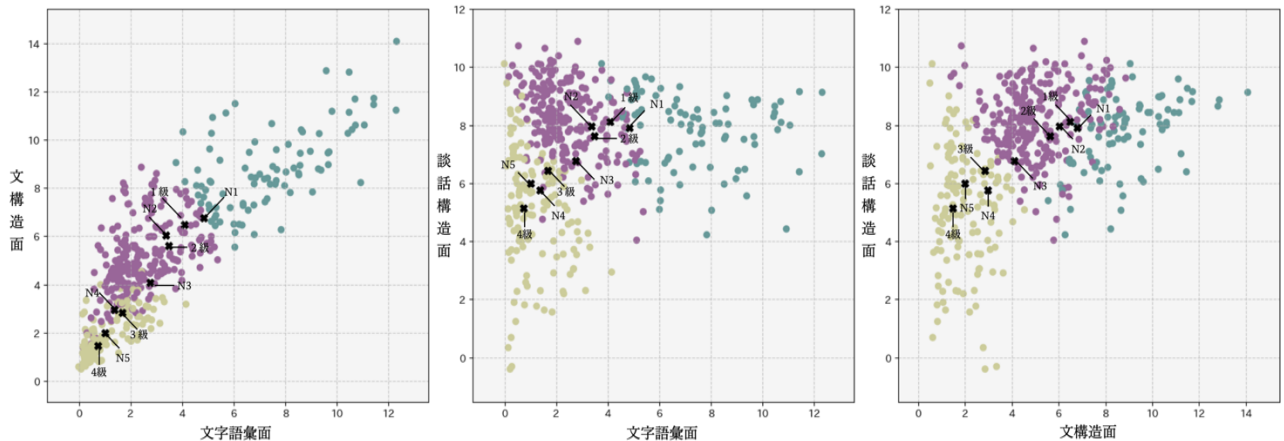


図 k-means法によるクラスタリングの結果

付録 C 文章例

どこか遠くへ行きたい、久しぶりに旅に出たいと思ったとき、あなたはどんな旅をしますか。鉄道で行くのんびりとした旅に出るのでしょうか。日本の各地にはいろいろな鉄道があり、とてもおもしろいです。海岸線や渓谷を走る鉄道、トロッキ列車や登山鉄道などその地でしか乗れない列車、ストップ列車やこたつ列車などその季節にしか走らない列車などがあります。

例えば、こんな旅はどうでしょうか。お薦めは静岡県の新金谷と千頭の間を走っているSLの旅です。SLは古いものでは、今から80年ほど前に製造されたものがあり、車内の天井には当時のままの鉄製の高風機が回っています。トンネルに入ると、石炭を燃やして出る煙が窓から入り込んでくるので、気を付けましょう。でも、それもおもしろさの一つです。SLに揺られながら静岡県産のお茶を飲み、駅弁を食べる気分は最高です。車窓に目を向けると、季節によって変わる山や川の景色がゆっくりと通り過ぎていきます。そして、車内では車掌さんがその土地の説明をしたり、ハーモニカで昔の歌を吹いたりして、乗客を楽しませてくれます。

鉄道で旅をしていると、ハプニングも起こります。乗り遅れたり、忘れ物をしたりと……。このような思いがけない出来事も全てが面白い思い出になります。通勤や通学の電車では乗客同士で話をするのはあまりありませんが、旅先では向かい合った席の人と自然に会話が始まることもあります。知らない人との会話もまた旅の楽しみの一つです。

旅をするときには、フリー切符が便利です。フリー切符は途中で列車を何度も乗り降りできるだけでなく、観光施設を割引で利用することもできます。駅や観光施設案内所で配布している観光マップやパンフレットも利用しましょう。ガイドブックには載っていない見どころやお店の情報を手に入れることもできます。そこに観光したときの感想などを書き込んでいけば、旅日記にもなります。

ときには日常を離れて、のんびりと鉄道の旅に出てみませんか。

付録 D LLM 判定用のプロンプト

教科書レベル判定のプロンプト

#指示
あなたは優れた日本語教師です。次の文章について、サンプルを参考に「初級」、「中級」、「上級」を判断し、その理由とともにJSON形式 (gradeとreasonがキー) で出力してください。

#対象の文章
{target}

JLPTレベル判定のプロンプト

#指示
あなたは優れた日本語教師です。次の文章について、日本語能力試験におけるレベル (N1 (上級), N2, N3, N4, N5 (初級)) を、「文字語彙」、「文構造」、「談話構造」の3つの観点から総合的に判定してください。それぞれの観点を10点満点で評価し、判定の理由とともにJSON形式 (grade, character_score, sentence_score, discourse_score, reasonがキー) で出力してください。

#対象の文章
{target}