

文章の書き分けに着目した言語運用能力の評価方法の検討
Development of Language Proficiency Evaluation Method Based on Classifying and Writing According to Targets

赤木 信也[†]
Shinya Akagi

納富 一宏[‡]
Kazuhiro Notomi

1. はじめに

学生や社会人にとって論文作成や報告書作成などにおける言語運用能力（表現力・語彙力）を高めることは重要である。本学では「文章表現」という講義において文章の体裁などに関する指導が行われているが、全ての学生に対する個別指導は教員の負担が大きい。

そこで本研究では、教員負担の削減を目的として、計算機による言語運用能力を評価・指導するシステムの開発を目指す。システムの開発では、「言語運用能力が高ければ、読者に合わせて文章の読みやすさを調整したり、分野に適した表現を用いたりできる」という考えに基づき、文章の書き分けに着目する。

本稿では、システム開発の前段階として、文章書き分けに着目した言語運用能力の評価方法を検討する。小学校高学年生向けの文章と大学生向けの文章の書き分け実験を行い、それぞれの文章について可読性分析を実施した結果について報告する。

2. 可読性指標

可読性指標は、文章の読みやすさ・分かりやすさを計算機に推定させるための基準のことである。英語に関してはFlesch Reading Ease^[1]が標準的指標として確立し、保険証書の読みやすさ保証のため、州保健法に数値基準が定められている。可読性指標の開発モデルには表1のような手法が存在する。

表1 可読性指標の開発モデル

線形式モデル	回帰分析、主成分分析による線形式を利用
語彙リストモデル	難易度毎に分類された語彙の割合を利用
言語モデル	難易度毎に分類された文書集合に対する最尤推定の結果を利用
変数置き換えモデル	英文に対する可読性指標、および日本語対応表を利用

「変数置き換えモデル」は我々が提案した手法であり、既存の英語指標の変数を同程度の値で日本語に対応付けることにより、サンプル文に依存せず、英日両文に適応可能な指標を開発する手法である。過去研究において、字種分割と漢字1字単位の解析、シャノン情報量に基づく重みづけを用いた日本語対応付けが良いという結果が得られている^[2]。本研究の可読性分析には、変数置き換えモデルに基づく指標として、jFRE, jFKG, jARI, jCLI, jSMOGを用いる。評価式を式1～式5に示す。

$$\begin{aligned} jFRE &= 206.835 - (1.015 \times ASL) - (84.6 \times ASW) && \dots\dots(式1) \\ jFKG &= (0.39 \times ASL) + (11.8 \times ASW) - 15.59 && \dots\dots(式2) \\ jARI &= (4.71 \times ACW) + (0.5 \times ASL) - 21.43 && \dots\dots(式3) \\ jCLI &= (5.88 \times ACW) - (29.6 / ASL) - 15.8 && \dots\dots(式4) \\ jSMOG &= 1.031\sqrt{30 \times PS} + 3.1291 && \dots\dots(式5) \end{aligned}$$

※ASL = 字種分割数/センテンス数
※ASW = 英語1音節・漢字1字単位の再分割数/字種分割数
※ACW = シャノン情報量に基づく重み/字種分割数
※PS = 英語3音節・漢字3字以上の字種分割数/センテンス数

ACWの日本語対応付けにおけるシャノン情報量に基づく重みは、数字・アルファベットを基準(1.000)として、文字種の種類数に応じて決定する。各文字種に対する重みを表2に示す。

表2 シャノン情報量に基づく重み

文字種	文字一覧	重み
数字 アルファベット	“0”～“9”, “a”～“z”	$-\log_2(1/36) / -\log_2(1/36) \approx 1.000$
平仮名	“あ”～“ん”	$-\log_2(1/83) / -\log_2(1/36) \approx 1.233$
片仮名	“ア”～“ヴ”, “ー” (長音)	$-\log_2(1/85) / -\log_2(1/36) \approx 1.240$
漢字	“一”～“龠”, “々”	$-\log_2(1/20898) / -\log_2(1/36) \approx 2.776$

本研究で用いる変数置き換えモデルに基づく指標は、評価基準表(表3参照)と照らし合わせることにより、文章を7段階の難易度で評価できる。本研究の可読性分析では、点数水準の評価としてjFREの評価値を用い、学年水準の評価としてjFKG, jARI, jCLI, jSMOGの平均値(Mean)を用いる。

表3 可読性指標の評価基準表

点数水準	学年水準	評価
90 - 100	5	Very Easy
80 - 90	6	Easy
70 - 80	7	Fairly Easy
60 - 70	8 - 9	Standard
50 - 60	10 - 12	Fairly Difficult
30 - 50	13 - 16	Difficult
0 - 30	17 and above	Very Difficult

3. 実験方法

本実験では、被験者に小学校高学年生向けの文章と大学生向けの文章の書き分けを行ってもらい、可読性分析を通じて、書き分け文章を言語運用能力(表現力・語彙力)の評価に利用できないか検討する。

[†] 神奈川工科大学大学院工学研究科
Graduate School of Engineering, Kanagawa Institute of Technology
[‡] 神奈川工科大学情報学部情報工学科
Kanagawa Institute of Technology

被験者は本学の納富・鈴木研究室に所属する計3名の大学生・大学院生を対象とする。文章の書き分けタスクとしては、「インターネットとは何か」という用語タスクと、「レポートのコピー&ペーストに関するあなたの考え」という意見タスクの2種類を用いる。各タスクの制限時間は10分とし、文字数のみを表示する「システム1」と、文字数に加えて点数水準と学年水準の可読性評価を表示する「システム2」を用いて文章の書き分けを行う。文字数は100文字以上なら○、100文字未満50文字以上なら△、50文字未満なら×を表示する。可読性評価は、小学校高学年向けの場合 Easy なら◎、その前後の評価なら○を表示し、大学生向けの場合 Difficult なら◎、その前後の評価なら○を表示する。

「システム1」を用いた実験では、用語タスクと意見タスクについて、特に指示を与えず、小学校高学年向けの文章と大学生向けの文章の書き分けを行ってもらった。

「システム2」を用いた実験では、用語タスクと意見タスクについて、可読性評価が◎または○になるように文章を調整するよう指示を与え、小学校高学年向けの文章と大学生向けの文章の書き分けを行ってもらった。被験者によるタスク最適化の影響を抑えるため、「システム1」を用いた実験から1週間以上の期間を空けて実験を行う。

4. 実験結果

「システム1」および「システム2」の書き分け文章について可読性評価の結果を表4、表5に示す。

表4 「システム1」の書き分け文章

	文字数	jFRE	Mean
平均値	134.2	78.6	7.7
標準偏差	24.3	9.2	2.2
中央値	132.5	75.9	7.9
最小値	105	68.1	4.0
最大値	187	99.5	11.5
◎		2	1
○	12	4	4
△	0	5	6
×	0	1	1

表4より、「システム1」では、jFREの評価記号は◎が17%、○が33%、△が42%、×が8%となっており、Meanの評価記号は◎が8%、○が34%、△が50%、×が8%となっている。5割程度が○以上の評価を得られている。

表5 「システム2」の書き分け文章

	文字数	jFRE	Mean
平均値	148.9	66.5	10.1
標準偏差	42.9	17.5	5.1
中央値	151	70.3	8.3
最小値	100	38.5	4.6
最大値	238	89.7	23.4
◎		6	2
○	12	3	5
△	0	3	4
×	0	0	1

表5より、「システム2」では、jFREの評価記号は◎が50%、○が25%、△が25%、×が0%となっており、Mean

の評価記号の割合は◎が17%、○が42%、△が33%、×が8%となっている。6割程度が○以上の評価を得られており、「システム2」における可読性評価の表示を受けて、被験者が文章を調整した結果だと考える。

次に、「システム1」および「システム2」の書き分け文章について可読性評価の比較を行った。用語タスク（小学校高学年向け、大学生向け）のMeanを図1、図2に示す。

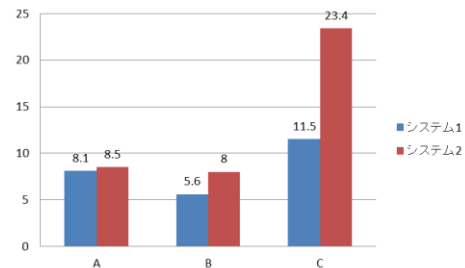


図1 用語タスク（小学校高学年向け）のMean

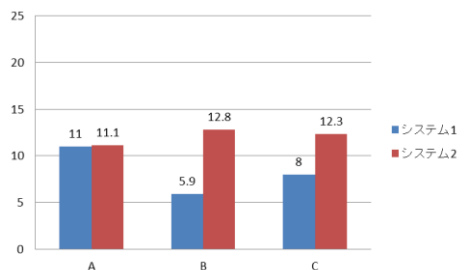


図2 用語タスク（大学生向け）のMean

図1ではMeanが6以上7未満が◎、図2ではMeanが13以上17未満が◎である。図1、図2より、被験者Aはシステム問わず書き分けができていた。被験者Bは「システム2」利用時に書き分けができていた。被験者Cは小学校高学年向けの書き分けが苦手であることが分かる。

最後に、実験後の被験者に対して書き分け実験および可読性評価に関する聞き取り調査を行った結果を示す。これらの指摘は今後の課題となる。

- ・語彙の調整として「書く」を「記述」にしたが、評価はそこまで変わらなかった。
- ・平仮名ばかりにすると易しいと判定されるが、逆に難しいのでは。
- ・「システム2」では計算機の評価に無理やり合わせたため、文章の良さはあまり関係なくなった気がする。完成した文書の評価に用いさせる工夫が必要ではないか。

5. まとめ

文章の書き分けと可読性分析を用いて言語運用能力の評価・支援ができないか検討を行った。書き分け文章の可読性評価に被験者間の差が生じており、これを言語運用能力の一つとして利用できる可能性が示唆された。

参考文献

[1] Flesch, R., "A new readability yardstick. Journal of applied psychology", 32(3), pp.221-233, (1948).
 [2] 赤木信也, 納富一宏, "字種分割を用いた日本語リーダビリティ判定システムの開発", 電気学会 電子・情報・システム部門大会講演論文集, pp.1192-1197, (2015).