

レクサイル指標の位置づけと計測方法

Methodology of Lexile Index and Its Foundation

藤原 郁郎
Ikuro FUJIWARA

In 1988 a private data mining firm, MetaMetric, created Lexile index. It is different from any other readability index in that it measures both students' comprehension capability and reading material difficulty to read. Since its publication, Lexile index has been prevailing in U.S. society to the extent that almost half school children are evaluated their reading comprehension capability with the index. Furthermore, Common Core State Standards, an education policy given by Federal Government, consequently promotes the use of Lexile index to improve students' reading capability. 150 publishers including McGraw-Hill and distributors such as amazon.com also facilitate the index in their business. The influence of Lexile Index is now covering U.S. education entities from K to 12. This paper discusses the general concept of readability index first. Then, it shows how to calculate a basic readability index. Finally, the methodology of Lexile Index is described and discussed.

KEY WORDS: Lexile Index Readability Index Quantitative Method
 MetaMetric Flesch Reading Ease

はじめに

英文の読みやすさを数値でもって示すことができるかどうかについては賛否両論がある。賛成論者はさまざまな英語文献をもとに構築されたリグレーション・モデルにより計測される数値には実用性があると主張をする (Chall and Dale 1949; McLaughlin 1964; Stenner et al. 2014)。反対論者によれば、計量言語学におけるわずか二つの要因に過ぎない語彙と文の長さのみを対象とした読みやすさの指標化は、文体、レトリック、比喩、含意、構造的複雑さなどの重要な要因が除外されているため大きな欠陥があると指摘する (Klare 1976; Armbruster et al. 1985; Bruce et al. 1986)。賛否両論があるなか、1948年に公表されたフレッシュ・リーディング・イーズ (FRE) は1970年代ごろから一部の政府文書で参考指標として使われ始めるようになった (Naval Technical Report Command 1975)。その後もリーダビリティ指標の利用は社会的な広がりを見せ、FREはMicrosoftワード2010に標準搭載され、法律や経営の文書作成においてある役割を果たすようになりつつあると言われている。たとえば新しく弁護士となった

人々が、上告の際の文書作成においてMicrosoftワード2010の読みやすさの指標を、従来の上告文書と同程度の読みやすさにするために活用されていることが報告されている (Long 2012)。

本稿においてはリーダビリティ指標の妥当性と信頼性は「ある限定的な条件」のもとでは高まることを示すとともに、指標の実用化における有効性の根拠として重視されなければならないことを論じた。これまでの研究者は指標化の際の少ない説明変数を批判し、実用性が低いことを具体例に抛りながら議論をしてきた (Armbruster et al. 1985; Bruce et al. 1986)。しかし限定的な条件を考慮することにより、特異性を有する文書に対して、計量言語学におけるいくつかの要因により指標化することには意義があると考えられる。

本稿では、最初に読みやすさの指標について基本的な理解のための特性について言及し、次いで問題点について論じた。さらにどのように計測されるのかについて実際の英文を用いながら方法論を示した。Microsoftワード2010に標準搭載されているFREを取り上げたが、FREはレクサイル指標が広がる以前、最も影響を与えた読みやすさの指標であった。最後に、レクサイル指標についても具体例を用いて計測方法を示した。ただし、メタメトリクス社が独自に構築した500万語にわたるコーパスは公表されていないため、筆者の推測部分が含まれており、その部分は明記した。

妥当性と信頼性について

読みやすさの指標化に対する理解を深めるためには、その妥当性と信頼性について議論をしなければならない。この場合の妥当性とは「読みやすさの指標」がさまざまな文書に対して実際に「読みやすさ」を指標化しているのかどうかを問うということである。語彙と文長という要因だけを説明変数としてリグレッション・モデルにより数量化された数値が、実際に「読みやすさ」という概念の指標として妥当なものかを問うことである。語彙と文長のみから計測された指標は、何かほかの概念、たとえば文章の「単純さ」のようなことを数量化したものと考えの方がより適切である、などと論じることなどである。妥当性の問題は、「読みやすさ」の指標が「読みやすさ」そのものを的確に示しているのかどうかを検証することである。

信頼性については、「読みやすさ」の指標として計測された数値が、実際に実用に耐えうるものかを問うことである。たとえば、筆者は本稿のためにヘミングウェイの“The Killers (1927)”のFlesch-Kincaid Grade Level (FK Grade Level) について読みやすさの学年指標を計測してみたが、その数値は小学1年生を示す1.1であった。この作品は、ヘミングウェイの初期のなかでも感情的な言葉を排除したいわゆるハードボイルドの作品であり、プロボクシングの八百長試合での裏切りにより生き方を見失ってゆくスウェーデン系ボクサーを描いている。「読みやすさ」の指標が小学1年生にも読むことができる数値であるからと言って、小学1年生の読書作品のなかに入れることはできない。「読みやすさ」の指標はコンピューターによりあらゆる文書

に対して自動的に計測されるわけであるが、この例で言えば、その数値の信頼性は低いと論じざるを得ないわけである。

妥当性や信頼性の問題を指摘したが、本稿においては「読みやすさ」の指標は妥当性と信頼性に問題があるものの、「ある限定された条件のもと」においてはこれらを実用化されるまでに高める可能性があるという視座のもとに議論を進めている。

読みやすさの指標の特性

数量的分析はコンピューター技術の進歩とともに、あらゆる学問領域へ深く入り込んでいる。計量言語学の分野においても数百万の語彙を有するコーパスが構築され、語彙のみならず作家による文体の特徴や比喩の統計的分析が行われている。「読みやすさ」の指標もこのような技術的な進歩に多くを負っていると指摘できるが、その基本的な特性を理解することは議論の土台として重要なことである。

第一に、読みやすさの指標は順序尺度であることを認識しなければならない。数量化された数値は、名義尺度、順序尺度、間隔尺度、比例尺度の4種類に分類することが一般的だ。比例尺度は長さ・広さなどに使われる数値であり、値を加減乗除しても意味がある。2倍の数値であれば実質的な変化が2倍あることを示す。しかし読みやすさの指標にはこのような比較は意味がない。10という数値を与えられた英文は、5という指標を与えられた英文よりも2倍の「むつかしさがある」ということではない。さらに読みやすさの指標は気温などの間隔尺度でもない。4種類の文書があったとして、一編ごとの読みやすさの指標が、85, 80, 45, 40と計測されたとする。85と80の差は5であり、45と40の差も5であるが、この差の数値5が示すものは「同じ読みやすさ（むつかしさ）の程度」を示しているわけではない。

読みやすさの指標は一般に「インデックス」という呼称が使われている。これは「索引」という意味がある。つまり、順序尺度としての数値が計測されているのではあるが、むしろ「読みやすい」、「読むのがむつかしい」というインデックス的な振り分けを細かく行ったものとも考えられる。この場合、順序尺度よりも名義尺度に近くなり、いわゆるレッテル貼りのようなものであるとも指摘できる。

したがって、読みやすさの指標は数値で表現されてはいるが、それらの数値は加減乗除しても意味はなさず、読みやすさの難易度を細かく分けていった順序尺度、ある場合には細かく分けられた名義的な尺度に近くなるということへの理解は、指標に対する過度の依存を避けるためにも必要なことと言わなければならない。

DuBay らによれば、読みやすさの指標は200を超えたものが発表をされているという（DuBay and Cheryl 2012）。それら個々の計測式から計測された数値を単純に比較することはできない。個々の計測式に使われるリグレッション・モデルの係数や定数項、さらにはそれらを決定する

ために用いた文書が異なっているためである。さまざまな読みやすさの指標の数値は異なった方法論のもとに計測されている。ただし、異なる計測式から得られた数値を参照しながら、分析対象の文書の読みやすさを考察することは有効なことと言わなければならない。その際、個々の計測式の意味を考慮し、その指標の範囲なども考慮することが求められる。実用で使われる場合には個別の利用者にとってはこのような方法論まで考察をした比較は煩雑となるため、単にいくつかの読みやすさ指標の値を参照するにとどまることが多くあると思われる。実際、学年別においては5つ程度の読みやすさの指標を算出し、その平均をとってより適切な学年を示すオン・ラインのウェブ・サービスもかなりある (e.g. readability-score.com)。このような場合、個々の計測式の方法論を抜きにして、結果のみを加減しているわけであり、その信頼性がある程度高まる場合もあるだろうが、結局、異なる計測式で算出された数値はその方法論において異なった意味があり、それらを加減することは大きな方法論的瑕疵が存することとなる。読みやすさの異なる指標に対しては、それらの「数値」を参照することにとどめるべきであり、加減をした数値は方法論にまで立ち至らなければならない別のことであることと認識しなくてはならない。

読みやすさの指標の特性として、順序尺度であること、各指標間の比較は方法論まで考察が必要であることを論じた。さらに言語的特性つまり英語とは異なる言語における読みやすさの指標との比較も行わなければならないが、本稿では言語的特性については言及しておらず、別に論じたい。

限定的条件 — 実用における留意点 —

読みやすさを数量化することにおいて妥当性と信頼性に大きな問題があることを指摘した。これらの問題をできる限り小さくすることが可能かどうかを議論することには意味がある。実際、既存の読みやすさの指標化においては、レトリカルな表現、比喩、含意、詩的表現を考慮しているわけではない。さらに作家の個性などについては、語彙の難易度だけではなく特徴的な語彙の繰り返しや文体における特性などを考慮しなければならず、読解の難易度とは異なる方法論が必要であることがわかる。結局、読みやすさの指標化は語彙の難易度と平均の文長により計測される「数値」であり、「文章の読みやすさ」のうちの側面のみにおいて妥当性があると指摘できる。

しかし、語彙以外の要素、たとえば比喩的表現についての理解についてみれば、子供たちの成長とともに進むものであり、使われた語彙の難易度で学年別に振り分けることも一つの方法論と考えることができる。「青空の大きな帽子をかぶって (wearing a large cap of blue sky)」という比喩的表現は小学生低学年でも理解できる。「歴史とは時間と空間が織りなす綾錦 (History is brocade woven with time and space)」という比喩的表現は中学生、「琴線に触れ

る話に慟哭を禁じ得ない (Heartstrings moved by a gale of apologue let loose its nature)」などの表現は高校生と、おおよそ心の発達とともに比喩的表現・レトリカルな表現への理解が進むものとして、数量化においてはこれらの表現が意味するものを評価するよりも、それに使われている語句の難易度を考えることで学年別に振り分けることはある程度意味があると考えることができる。

現在、アメリカの実社会で使われている読みやすさの数量化は、「学年別」の指標を与えるものが多くある (e.g. Flesch-Kincaid Grade Level and Gunning Fog Index)。「学年別」に設定することにより、読者を限定させ、専門性の問題や読み手の能力の多様性の問題を最小限に抑え、妥当性を高めることがその背景にあると指摘できる。

「学年別」ではない指標の場合にも、その使用方法を工夫すれば妥当性と信頼性を高めることができる可能性がある。これらの指標の代表は先にも述べたワード2010にも標準搭載されているフレッシュ・リーディング・イーズ (FRE、Flesch Reading Ease) だ。FREは連邦政府にも採用されており、米軍の新兵たちが読む文書に適用され軍の通達がより正確に行き届くための一助として使われている。米軍の新兵たちが読む軍のレポートでは、18歳前後の若者が読む軍の勤務規定や初歩のテクニカル・レポートが中心であり、その文書を作成する書き手についても軍での経験がある程度あり、軍の用語や表現形式にも習熟していることが予想される。詩的でレトリカルな表現などはまったく必要がないことは言うまでもない。このような例では、(1)文書の一文の平均単語数、つまり長い文が多ければ多いほど読みづらさは増し、(2)音節数が多い単語が多ければ多いほど文書は読みづらくなるだろう、と推測することは意味があると考えてよいだろう。実際、FREでは、この二つの要因を説明変数とし、多くの英文から(1)と(2)を計測して、読みやすさを0～100の数値として示している。100に近いほど読みやすいものとなるように係数と定数をあてはめてリグレッション・モデルを構築することにより数値化を行っている。

指標の背景 — 移民国家アメリカと英語教育 —

ある条件のもとでは、文章の「読みやすさ」を数量化することは妥当性と信頼性を高めることができることを指摘した。この前提は英語・日本語に限らず言語一般に当てはまる普遍的な性質として考えられる。しかし、読みやすさの研究は、1920年代、アメリカにおいて本格的に始まっている。これにはアメリカの政治的・経済的・歴史的・文化的な要因が深くかかわっていたと考えられる。1900年初頭から始まった「進歩（改革）の時代」は、アメリカの教育界にも大きな変化を与えた (Lagemann 1997)。変化の潮流の一つは、デューイーのプラグマティズムの哲学を背景とする生徒を中心とした「民主的教育」、今一つはエドワード・ソーンダイクを中心とする科学的教育論をもとにした「マス教育」であった (Levin 2012)。教育にかかわる

州政府はデューイの方法論に力を注ぐ時間的・人間的な余裕はなく、ソーンダイクの「マス教育」における科学的方法論を重視する方針を採用し、「標準化テスト」が開発され、初等・中等・高等教育の実際の教育の場で使われるようになる (Shavelson 2007)。第一次世界大戦において自国が戦場となることなく、戦後、世界の経済力と国際政治のパワーを獲得したアメリカへ、中・東欧からの大量の移民が押し寄せたことは、「マス教育」をいかに効率よく行うかの問題を突き付けたことになる。新しい移民の工場労働者たちはしばしば母国語ではない英語の読み書きに困難をきたしたため、YMCAをはじめ英語の読み書きのために開設された成人教室を訪れることも珍しくはなかった (Korman 1965)。英語を話さない大量の移民とその子弟に対して、どれほどの英語運用能力をどの学年までに習得させるべきかが、社会の要請と相まって切実な問題となったと考えられる。学年別に英語の読み物をより正確に配当する必要性が高まっていた。移民の子弟が相対的に多いクラスではやさしい読み物が選ばれ、少ないクラスでは高い難易度の読み物が選ばれることは自然なことであり、統一的な教育の質を保つことは困難となっていた可能性がある。

アメリカの「読みやすさ」の指標化には、移民国家アメリカの教育カリキュラムにおける切実な要求があったと言って良い。このような背景がすべての国にあるわけではないが、グローバル化とグローカル化が進展する現在の世界において、移民を多く受け入れる国もあれば、英語を第二公用語的に使用する国々も多くなりつつある。これらの国々では子供たちの自国言語と英語の読み物について、学年別配当の的確な読み物の配分の要請はこれまでになく高まっていると考えられる。この意味において、各国における「読みやすさ」の数量化研究は、今後とも発展し続けることが予想される。

指標の計測方法 — FRE を例に —

計量的・数量的研究の方法論に関する研究は19世紀後半に教育学・心理学・生物学の分野で盛んとなった。教育学では生徒のテストにおける「真の学力」と「測定誤差」と言う基本的な方法論は現在も使われ続けている (Dietel et al. 1991)。全数調査から標本調査のサンプリング理論もこれらの分野でまず発展した。また心理・教育学者により作られたIQ (Intelligent Quotient、知能指数) は多くの批判があるものの早くも1905年に発表されている。

レクサイル指標の計測方法では、統計学的手法のうちリグレーション・モデルを用いている。リグレーション・モデルは現在、経済学や医学など「予測」が必要な分野で幅広く活用されている。その要点は次のように示される。

$$Y (\text{読みやすさの指標}) = (-1.015) \times X_1 + (-84.6) \times X_2 + 206.835$$

$$X_1 = \text{一文の平均単語数 (一文の長さの平均)}$$

$$X_2 = \text{一単語の平均音節数（単語の音節としての長さの平均）}$$

ここに挙げたリグレーション・モデルはフレッシュが最初に発表をした読みやすさの指標であるフレッシュ・リーディング・イーズの計算式である（Flesch 1948）。この式を見てわかる通り、ある英文の一文の単語数を数え上げてその平均をとり、総音節数と単語数を数え上げて割ることにより一単語あたりの平均音節数が計算され、この式に当てはめればFREが計算できることになる。実際に国内の英字新聞の記事を使ってFREを計測してみると次のようになる。

With Japan set to leave for the Rugby World Cup on Monday, local fans got a taste of the history of the sports this week with the first-ever visit to these shores by Rubgy School. Back in 1823 a 16-year-old schoolboy at the English private school “first took the ball in his arms and ran with it.” As a plaque commemorating the event states, William Webb Ellis did so “with a fine disregard for the rules of football as played in his time.” Little did he or his startled teammates know what the consequences would be.

文の数 4, 全単語数 104, 総音節数137

英字新聞からとったこのパラグラフの文の数は4、全単語数は104であるので、一文当たりの平均単語数は26語となる。また1音節語も含めた総音節数は137であるので、一単語あたりの平均音節数は全単語数で割った1.36となる。この26と1.36をそれぞれ X_1 と X_2 に代入をしてFREを計算すると次の値が得られる。

$$\begin{aligned} Y (\text{読みやすさ指標 : FRE}) &= (-1.015) \times 26 + (-84.6) \times 1.36 + 206.835 = \\ &= 65.4 \end{aligned}$$

FREの数値が60～70は、アメリカの13歳～15歳の生徒が理解できる難易度を示している。

ただし、計数の-1.015と-84.6と定数項の206.835をどのように決めているのかについては、さまざまな英文資料のもとに計算されている。これらの値は、Yの値が0～100の間で100に近いほどやさしいように決められている。専門書などの場合にはほぼ30以下となるように、また、小学生の読み物であれば60以上の数値となるように決めてゆくわけだ。「決めてゆく」と一言で表現をしているが、実際には膨大なマトリックス（行列）を処理してゆかなければならない。一例として筆者が作成したYと X_1 、 X_2 との関係を示した表を掲げる。

X ₁ : 文長	X ₂ : 平均音節数									
	1.00	1.15	1.30	1.45	1.60	1.75	1.90	2.05	2.20	
28.0	93.82	81.13	68.44	55.75	43.06	30.37	17.68	4.99	-7.70	
25.5	96.35	83.66	70.97	58.28	45.59	32.90	20.21	7.52	-5.17	
23.0	98.89	86.20	73.51	60.82	48.13	35.44	22.75	10.06	-2.63	
20.5	101.43	88.74	76.05	63.36	50.67	37.98	25.29	12.60	-0.09	
18.0	103.97	91.28	78.59	65.90	53.21	40.52	27.83	15.14	2.44	
15.5	106.50	93.81	81.12	68.43	55.74	43.05	30.36	17.67	4.98	
13.0	109.04	96.35	83.66	70.97	58.28	45.59	32.90	20.21	7.52	
10.5	111.58	98.89	86.20	73.51	60.82	48.13	35.44	22.75	10.06	
8.0	114.12	101.43	88.74	76.05	63.36	50.67	37.98	25.29	12.60	
5.5	116.65	103.96	91.27	78.58	65.89	53.20	40.51	27.82	15.13	
3.0	119.19	106.50	93.81	81.12	68.43	55.74	43.05	30.36	17.67	

(資料作成：藤原)

一文の長さが短くなればなるほど、一単語の音節数が少なくなればなるほど読みやすさの指標が大きな値となっていることが示されている。実際にはデータに最もよく合うように最小二乗法や最尤法をつかった計算を行わなければならない。この計算は現在ではExcelをはじめ統計処理用のアプリケーション・ソフトで容易に行うことができる。筆者は、FREの計算式において係数を変える必要があると考えている。その理由はフレッシュがFREを発表した1948年当時と現在では、一文の長さがさらに短く、単語の音節数も少なくなっている可能性があるからだ。この問題については別に論じなければならず、本稿ではこの点の指摘にとどめておきたい。

レクサイル指標 — 読み手と読み物の数量化 —

各国の子供たちが、本国言語の「読みやすさ」の数量化により、より適切な読み物が与えられる可能性について言及したが、この限定条件がレクサイル指標の背景となっている。レクサイル指標とは小学校・中学校・高校生が、自分の読解能力に見合った読み物の指標として示されているものだ。したがって、英文一般の読みやすさの数量化を目的としたものと考えるよりも、英語運用能力の教育課程における英文の読みやすさと子供たちの読解能力を示す指標として開発されたと考えるとより適切な理解へとつながる。

なぜ、レクサイル指標が教育現場で使用されるようになったのかについては、明らかな理由がある。FREや従来の「読みやすさの指標」では、読み物に対してのみ数量化された数値が与えられていたが、レクサイル指標では、読み物とともに個々の生徒の読解能力にも指標が与えられているからだ。レクサイル指標そのものは一義的な公式から計算される数値であるが、その数値は「読み物」とともに個々の「生徒の読解能力」にもあてはめられている。

しかし、子供の読解能力は書物へのアクセスのしやすさに大きく影響をされている報告もな

されている（Neuman and Celano 2001）。またレクサイル指標の計測は計量言語学的には語彙と文長の二つの説明変数のみを活用するものであるため、従来のテキストの複雑性の問題が依然として指摘されている（Kraschen 2002）。

アメリカ教育省は、レクサイル指標の公立学校での利用が進む状況を受けて、ブラウン大学のホワイト教授のもと諮問委員会を設置し、レクサイル指標の教育の場での使用について答申を求めた。ホワイトらは複雑性の問題、語彙に対する分析が限られていることなどを指摘するとともに、数値化される読解能力の心理的影響まで言及し、レクサイル指標の教育の場での使用は限定的に行うべきであると結論した（Ministry of Education 2001）。

現在もレクサイル指標の有効性については議論が続いているものの、初等・中等教育におけるテキストの難易度の比較においては、計量分析的手法の「読みやすさの指標」に依拠しなければならず、Common Core State Standardsにおける読解能力の指針において、Lexile Frameworkが使われるようになったことはある意味、順当なことである。質的分析においてはあらゆる人々が数千冊の書物について難易度について一致した見解を見出すことは困難であり、より客観的な議論のためには量的分析に頼らざるを得ない側面を示している。こういった背景のもと、連邦政府が2010年に各州共通の教育基盤指針（コモン・コア）の概要が打ち出されるに及んで、レクサイル指標は全米の初等中等教育で広く使われるようになっていった。私企業により開発された読みやすさの指標が、今や連邦政府の教育指針と深くかかわりながら全米に拡大をし、その勢いはアメリカを超えて世界的にも影響を与えている。

次にあげるレクサイル指標は、メタメトリクス社の公式ウェブサイトに掲載されている資料に拠った（MetaMetrics web site 2015）。たとえばレクサイル指標が750Lのある読み物Aがあったとする（LはLexile indexを示す）。レクサイル指標が740L～940Lの読み物は小学4年生が読むテキストとして適切であることがあらかじめ示されているので、750Lは4年生配当の読み物として適切とされる。ところで、ある小学4年生Xが英語の読解テストを受験することで750Lというレクサイル指標が与えられとする。この場合、この生徒Xは、750Lの指標を持つ読み物Aの75%以上を理解することが可能である、ということを示している。さらに別の読み物Bのレクサイル指標が1000Lであったとすれば、この生徒XはBの約50%の内容を理解できる可能性が統計学的に有意になることを示している。レクサイル指標がおおよそ250L離れば、内容の理解度がおおよそ25%上下するという。したがって、この生徒がレクサイル指標500Lの読み物を読んだ場合には、おおよそ90～100%の理解度が予想されることになる。

レクサイル指標は、190L～530Lの小学1年生から1185L～1385Lまでの高校2・3年生までを対象としているが、各学年ともに数値に重複部分があり、幅広い運用が行われるように配慮されている。0L以下の絵本などについてはビギニング・リーダー（BR）の数値が設けられている。0Lから易しくなるにつれて大きな数値が与えられ、現在最もやさしいとされる絵本の数値はBR-420Lである。BRはBeginning Readers、LはLexile indexを示しており、マイナス記

号がつくのが本来の使い方が省略されている場合もある。

メタメトリクス社によるとレクサイル指標が与えられている書籍は2014年現在13万5千冊ほどだが、新聞記事やウェブ・サイトを含めると総数は1億部を超えており、オンラインの資料検索サービスであるProQuestではほぼすべての記事・資料に対してレクサイル指標が示されているという (MetaMetrics official web site 2015)。また先にも示した通り全米50州にわたる計3,500万人の生徒がレクサイル指標を与えられ、世界180か国でレクサイル指標が何らかの形で使われているという。日本でもamazon.co.jpはレクサイル指標を子供たちの英語読み物の指標として積極的に取り上げている。また「はじめに」において示したように、いわゆる「読み・書き・そろばん」を中心的な内容とする各州共通基盤スタンダード (CCSS) において、読書における読みやすさ指標の位置づけの利用が拡大され続けている。今後ともレクサイル指標は教育の場でますます活用されてゆく可能性が高いと言わなければならない。

レクサイル指標の計測方法論

レクサイル指標は、一般のリーダビリティ指標と同様、(1)一文の平均単語数、(2)出てくる単語の難しさ、という二つの要素で計算される (Stenner 1996)。ただし、(2)の「単語の難しさ」については、FREの計測方法で示したような「読みやすさの指標」で多く使われていた「単語の平均音節数や文字数」を使わずに、メタメトリクス社が独自に開発をした500万語に上るコーパスを活用し、頻度数を単語の難易度として処理されている。一般の文書に多く出てくればそれだけ難易度は低いと考える考え方は、1921年に「教師用マニュアル1000語」を完成したソーンダイクの考え方と同じものである。

ソーンダイクからレクサイル指標までの間に発表された数量化式では、単語が長ければ長いほど難易度が上がるという原理に拠るものであるが、エリザベータンから現代までの文書を比較したシャーマンの先駆的研究に端的に示されているように、書き言葉は口語化へと向かってきた (Sherman 1893)。したがってシラブル数の多い単語が多く出てくる文書ほど読解がむづかしくなる、という考え方は言語学的な妥当性を有していると言わなければならない。しかし、「読みやすさ」の指標化を最初に試みたとされるソーンダイクがなぜ単語の出現度のリスト作成にこだわったのかを振り返れば、単に単語の長さを難易度の指標とするよりも、単語がよく使われるものかどうかを基準にすることにより妥当性はより高まると考えられる。1921年に「教師用マニュアル1000語」を完成したソーンダイクの方法論が、1988年に至って、単語リストのレベル分け方式が復活をしたことは、ソーンダイクの見聞性がコンピューター技術の進展とともに実現化されたものとしてたいへんに興味深いことである。

次にレクサイル指標について解説するが、方法論はメタメトリクス社の開発者のステナーの論文に拠っている (Stenner 1996; Stenner and Budick 1997)。

レクサイル指標の位置づけと計測方法（藤原）

レクサイル指標は、上記の(1)と(2)の要因において、(1)は従来と変わらず一文当たりの単語数（一文の長さ）を用いているが、そのログの値をとっている。また(2)については、分析する英文全体のすべての単語一つ一つの出現頻度のログの値の平均をとっている。

[1] 一文当たりの平均単語数のログの値

[2] 全文に現れる単語をコーパスにより出現頻度を計測しログの値をとり平均した値

[1] と [2] の数値から次の式でロジット・スコアを算出する。

[3] ロジット・スコア = $(9.83247 \times [1]) - (2.14634 \times [2]) - C$ （Cは定数 constant）

ここでログとは対数（ロガリズム）のことであり、 $\log_{10}1000 = 3$ というように底10の場合には10の指数部分はその値となるものである。またロジットとは、 $\text{logit} = \log(M/(1-M))$ とある数値のオッズ比（それ自身を1からそれ自身で引いた数値で割ったもの）のログの値である。最後に [3] の logit score を用いてレクサイル指標を次の式から計算する。

[4] レクサイル指標 = $([3] + C) \times CF + 200$ （CFは conversion factor 変換係数）

開発者のステナーは定数CとCFを次の連立方程式の解としている。

$$(-3.3 + C) \times CF + 20 = 200$$

$$(2.26 + C) \times CF + 200 = 1200$$

理論的背景については込み入った議論が必要なのだが、要するにレクサイル指標の値をおおよそ200~1200程度に抑えるための定数であり変換係数である。開発者のステナーは一例としてこの連立方程式をあげている。この例でCとCFを計算するとC=3.3、CF=180となる。したがって、レクサイル指標の算出式にこれらを代入すると次の式が得られる。

[4]' レクサイル指標 = $([3] + 3.3) \times 180 + 200$

それでは実際にある例文を示しながら、レクサイル指標を求めてみる。

I went to Tokyo so that I could see my girlfriend. Those days, she worked for the ontological medical clinic.

(1)のログの値については、2つの文から成りそれぞれ11単語と9単語で構成されているので、一文当たりの平均単語数は10となり、 $\log_{10}10 = 1$ であるので [1] の数値は1となる。また、(2)のロジットの値では、メタメトリクス社のコーパスは公表されていないので、Google Ngramを使ったものを代用して、おおよその値を計算する。

単語	頻度(%)	ログ(log)の値
1. I	0.4215	3.6248
2. went	0.0234	2.3692
3. to	2.2451	4.3512
4. Tokyo	0.0015	1.1761
5. so	0.1457	3.1635
6. that	0.8234	3.9156
7. could	0.0821	2.9143
8. see	0.0634	2.9143
9. my	0.0978	2.9903
10. girlfriend	0.0003	0.47712
11. those	0.0782	2.89321
12. days	0.0249	2.39620
13. she	0.0906	2.95713
14. worked	0.0072	1.85733
15. for	0.6459	3.81017
16. the	5.1327	0.71035
17. ontological	0.0002	0.30103
18. medical	0.0067	1.82608
19. clinic	0.0006	0.77815

以上の出現頻度のログの値をとりその平均を計算すると-1.18233となる。ここで、対数の平均をとる意味を考える。対数の平均値は幾何平均と呼ばれるものであるが、一般社会の算術平均とは大きく異なり、幾何平均とはデータの値をすべて掛け合わせて、そのデータの数だけのべき乗根のことである。これは値が大きく異なる場合に用いられる。単語の出現頻度は、1万や10万単位の違いがあることも稀ではないため、通常の算術平均では出現頻度の非常に高いものや非常に少ないものにデータの平均が大きく偏ってしまう。それを補正するために幾何平均が使われているのだ。3つのデータを使って幾何平均の算出方法を簡潔に示す。

$$\text{算術平均} = (X_1 + X_2 + X_3) / 3 \quad \text{幾何平均} = (X_1 \times X_2 \times X_3) \text{の3乗根}$$

上述の19個の単語のログの値の平均値が-1.18233であるが、これを開発者のステナーはロジット・スコアと読んでいた。このロジット・スコア [3] を用いて、レクサイル指標は次の式で計算される。

$$\text{レクサイル指標}(L) = ((-1.18223 + 3.3) \times 180) + 200 = 581L$$

レクサイル指標の位置づけと計測方法（藤原）

上記の英文はレクサイル指標が581Lであることが示された。各学年のレクサイル指標は

小学1年生	190L～530L	中学1年生	970L～1120L
小学2年生	420L～650L	中学2年生	1010L～1185L
小学3年生	520L～820L	中学3年生	1050L～1260
小学4年生	740L～940L	高校1年生	1080L～1335L
小学5年生	830L～1010L	高校2・3年生	1185L～1385L
小学6年生	925L～1070L		

と示されている（MetaMetrics official web site 2015）ので、例文は小学2年生から小学3年生で75%以上の理解ができるであろう、ということを示している。実際には ontological などの難易度の高い単語が出現しているため、小学低学年では理解はできないのであるが、この英文の75%以上を理解する、というレクサイル指標を持つ意味にはおおよそ当てはまると考えてよいだろう。

おわりに

レクサイル指標は、コモン・コア・ステート・スタンダードのもとでさらに活用が進んでゆくと考えられるが、わが国の英語教育においてもその活用の必要性に迫られるのは必至の情勢である。英文の読みやすさを指標化することは絶対的な数量化は不可能ではあるが、限定された条件のもとでは、数量化の妥当性と信頼性が高まるということから議論を始めなければならないだろう。英語教育に携わるすべての人々が、数量化に対する理解と批判的考察を深めてゆくことが求められる。

読みやすさの指標という概念はわが国では発展をしなかったが、その理由は英語を使う人々が一部に限られていたこと、また日本語の特性が英語とは大きく異なることなどの理由が考えられる。しかしながら、英語教育が小学校3年生の正課となることが決定されている現在、英文の読み物の学年配当などの問題は必ず起こってくることであろう。その際、単にアメリカの教育で使われている指標化された読み物に準じて配当をすればよいのか、それとも外国語としての英語に対してなんらかの日本的な条件を考慮しながら、より内容面での配慮をするべきなのか。問題は山積みしていると言えるだろう。

しかし、レクサイル指標に対する理解が深まり、読みやすさの指標化への批判的考察や議論が高まれば、わが国における英語低年齢層化に伴う、英語読解教材に対する一つの確固とした哲学の形成にまで至るものと期待される。

「数値」を単なる数値ではなく、それが計測される背景への理解を深めながら、レクサイル指標や実用されている各種の読みやすさの指標の批判的運用が行われなければならない。

References

- Armbruster, Bonnie B., Jean H. Osborn, and Alice L. Davison. (2012). Readability Formulas May be Dangerous to Your Textbooks. *Educational Leadership*, 42(7), 18-20.
- Bruce, Bertram C., Ann Rubin, and Kathleen C. Starr. (1981). "Why readability formulas fail." Reading Education Report No.28 (1981, August). Urbana, IL: University of Illinois, Center for the Study of Reading
- Common Core State Standards Initiative. (2015). Official Web Site. <http://www.corestandards.org/>
- Dale, Edgar, and Jeanne Chall, (1949) The Concept of Readability. In *Readability* ed. by Edgar Dale. National Council of Teachers of English: pp.1-7.
- Dale, Edgar, and Jeanne S. Chall. (1949). A Formula for Predicting Readability. *Educational Research Bulletin*, 27(1), 11-28.
- Davison, Alice and Robert N. Kantor. (1982) On the Failure of Readability Formulas to Define Readable Texts: A Case Study from Adaptations. *Reading Research Quarterly*, 17(2), 187-209.
- Dietel, R. J., J. L. Herman, and R. A. Knuth. (1991). What Does Research Say About Assessment? *North Central Regional Educational Laboratory*.
- Dubay, William H. (2007). *Unlocking Language*. New York: BookSurge Publishing.
- Flesch, Rudolf. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, 32(3): 221-233.
- Ginges, A. M. (1958). The Readability Story. *The Australian Quarterly*, 30(4), 87-93.
- Klare, George R. (1975). Assessing Readability. *Reading Research Quarterly*, 10(1), 62-102.
- Klare, George R. (1976). A Second Look At The Validity of Readability Formula. *Journal of Reading Behavior*, 8(2), 129-152.
- Korman, Gerd. (1965). Americanization at the Factory Gate. *Industrial and Labor Relations Review*, 18(3), 396-419.
- Langemann, Ellen Condliffe. (1997). Contested Terrain: A History of Education Research in the United States, 1890-1990. *Educational Researcher*, 26(9), 5-17.
- Levin, R. A. (1991). The debate over schooling: Influences of Dewey and Thorndike. *Childhood Education*. 68(2), 71-75.
- Long, Lance. (2012). Does the Readability of Your Brief Affect Your Chance of Winning an Appeal? — An Analysis of Readability in Appellate Briefs and Its Correlation with Success on Appeal. *Stetson University College of Law Legal Studies Research Paper Series Research Paper No.2010-10*.
- McLaughlin, G. Harry. (1969). SMOG Grading – a New Readability Formula. *Journal of Reading*, 12(8), 639-646.
- MetaMetrics (2015) Official Web Site: <https://www.metametricsinc.com/>
- Neuman, Susan B. and Donna Celano. (2001). "Access to Print in Low-Income and Middle-Income Communities: An Ecological Study of Four Neighborhoods." *Reading Research Quarterly*, 36(1), 8-26.
- Shavelson, Richard J. (2007). A Brief History of Student Learning Assessment. *A Series of Academic the Association of American Colleges and Universities*. AAC&U.
- Sherman, Luciano Adolno. (1893). *Analytics of Literature*. Boston: Ginn and Company.
- Smith, Ron F. and Kay L. Smith. (1984). A Comparison of Readability Formulae as Applied to Newspaper Stories. *Journal of Reading*, 28(1), 20-23.
- Stenner, A. Jackson. (1996). Measuring Reading Comprehension with the Lexile Framework. Paper Presented at the Fourth North American Conference on Adolescent/Adult Literacy February

1996.

Stenner, A. Jackson, and Donald S. Burdick. (1997). The Objective Measurement of Reading Comprehension In Response to Technical Questions Raised by the California Department of Education Technical Study Group published by U.S. Department of Education.

White, Sheida, and John Clement. (2001). *Assessing the Lexile Framework: Results of a Panel Meeting Working Paper No.2001-08* U.S. Department of Education. Office of Educational Research and Improvement.

U.S. Naval Office. (1975). Derivation of New Readability Formulas For Navy Enlisted Personnel. Naval Technical Training Command Research Branch Report: pp.8-75.