



青山学院大学文学部紀要第 51 号 (2009: 97-120)

オンライン版「英文語彙難易度解析プログラム」(Word Level Checker) の概要およびその教育研究分野での応用可能性

染谷 泰正

Word Level Checker — An Online Lexical Profiling Program and its Pedagogical Applications

Yasumasa SOMEYA

1. はじめに

筆者の所属する青山学院大学文学部英米文学科では、主として同文学部で開講されている「英文講読」の授業に使う教材の難易度を客観的に測定するための支援ツールとして表題のプログラム（以下、WLC）を開発し、2006 年度からこれを英米文学科のウェブサイト上で試験的に公開している。このオンラインツールは、英米文学科内に設置されている「英語科目検討委員会」の要請によって作成したもので、たまたま筆者がすでに AWK という簡易言語を使って作成してあったプログラム^[1]を、改めて Java による CGI プログラムに書き換えたものである。

WLC については、ウェブサイト上での公開後もその概要について一般に紹介する機会を得なかったが、2008 年度になって筆者に時間の余裕ができたこと^[2]、および当初からの懸案事項のひとつであった「リーダビリティ」（テキストの「読みやすさ」の指標）の自動判定機能の追加作業が完了したことから、本稿を借りて改めて現時点での WLC の概要を紹介し、併せてその教育研究分野での応用可能性および今後の課題等について議論する。なお、従来の WLC の機能については別途オンラインの解説が用意されていることから、本稿では新たに追加されたリーダビリティ評価機能を中心に解説する。

2. WLC の概要

前述のとおり、WLC は主として筆者の勤務校で「英文講読」をはじめとする英語科目を担当しておられる専任および非常勤の先生方に使っていただくことを想定したものであるが、プログラム自体は英米文学科のウェブサイト上で公開しており、一般の方でも自由に使えるようになっている（アクセス先は <http://www.cl.aoyama.ac.jp/english/newSite/wlc/>）。プログラムの概要はおよそ次のとおりである。

2.1 初期画面：ユーザインタフェースの構成と使用法

上記のサイトにアクセスすると、図 1 のようなユーザインタフェースが表示される。図 1 に示したのは英文版であるが、日本語版も用意されている（画面左上の Japanese というリンクをクリック）。冒頭のタイトルの下には簡単な解説があり、その末尾にはより詳しい解説ページ

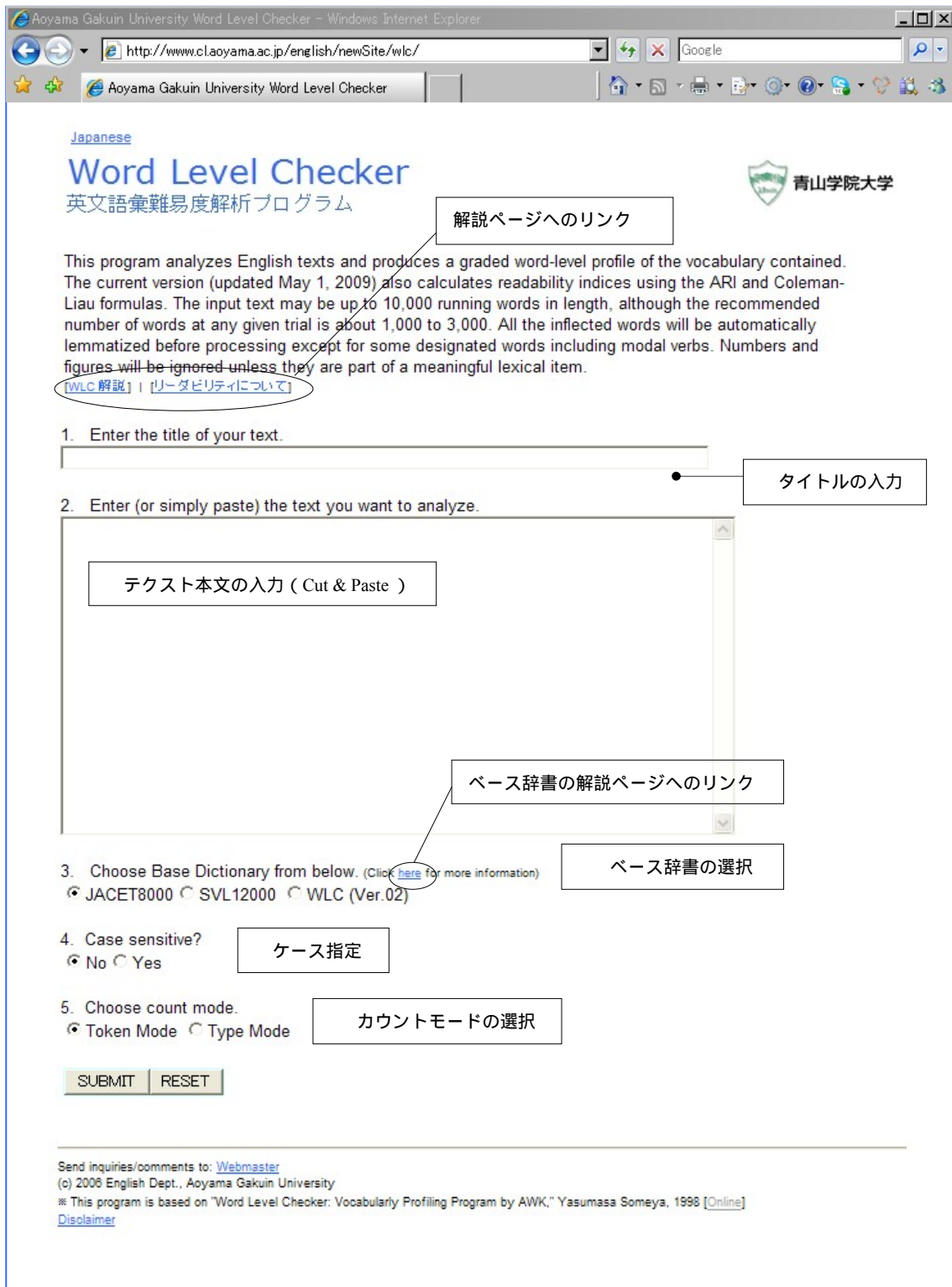


図 1 Word Level Checker (Online Version) の初期画面

画面は一部簡略化してあります。

へのリンクが貼られている。ユーザはまず最初の入力ボックスに分析対象のテキストのタイトルを入力する(ステップ)。タイトルの入力はオプションであるが、複数のテキストを分析

する場合、どの結果がどのテキストに対応したものがわからなくなることがあるため、テキストごとに適宜タイトルを加えておくのがよい。タイトルの次は、テキスト本文を入力する（ステップ ）。ボックス内に入力可能なテキストの量は基本的には制限は設けていないが、ストレスなく処理できるテキスト量として、およそ 3000 語から最大 1 万語程度を推奨している。なお、テキストはキーボードから直接入力してもかまわないが、入力ミスなどを防ぐため、あらかじめテキストファイルまたは MS Word ファイルとして作成したものを入力ボックス内に切り貼り (Cut & Paste) するのが望ましい。

ステップ では難易度解析に使用するベース辞書 (Base Dictionary) を指定する。現在のシステムにはベース辞書として JACET 8000, SVL 12000, WLC (Ver. 02) の 3 つのオプションが用意されている。JACET 8000 は大学英語教育学会基本語改訂委員会が編纂した学習語彙リストで、従来の基礎 4000 語 (JACET 4000 = 高校までの学習語彙約 3000 語 + 大学教養課程で習得が期待される 1000 語) を大幅に拡張した基本語彙 8000 語が収録されている。SVL 12000 は株式会社アルクが独自に開発した語彙リスト (「標準語彙水準 12000」) で、JACET 8000 等の既存の学習語彙表ではカバーされていない上級 ~ 最上級レベルの語彙 (Level 8 ~ Level 12) を含む 1 万 2000 語が収録されている。WLC (Ver. 02) は染谷 (1998) において作成されたもので、ビジネス英語の分析用に特化した約 3 万 5000 語の見出し語からなる語彙リストである。このうち JACET 8000 については大学英語教育学会から、SVL 12000 については株式会社アルクからそれぞれ使用許諾をいただいた^[3]。なお、各ベース辞書 (語彙リスト) の詳細についてはオンラインの解説を参照していただきたい。

ステップ では必要に応じて大文字・小文字の区別を指定する。初期設定値は No (区別しない) で、これを選択した場合、入力テキストの単語はすべて小文字化された上で処理される (例えば We we, Go go, Will will, Japan japan, JAPAN japan 等)。何らかの理由で両者の区別をしたい場合は Yes (区別する) を選択する。

最後に、ステップ で語数カウントの方法 (Count Mode) を指定する。Token Mode は出てきた単語を (同じものであっても) すべてひとつずつ数える。例えば、go という単語が go, goes, going, go, went, go の順に使われていた場合、変化形がすべて基底語化^[4]された上で、go = 6 (go が 6 回出現) と判定される。一方、Type Mode では基底語ベースで同一の単語は何回出てきても 1 回と数える。したがって上記の例では go = 1 ということになる。この選択はテキストの難易度判定に大きく影響するため、目的に応じた適切な選択をする必要がある。

以上の指定・選択が終わったら、画面下部にある [SUBMIT] をクリックする ([RESET] をクリックするとすべての指定・選択が解除されて初期画面に戻る)。通常、結果は数秒で返されてくるが、サーバや回線の状態によってはこれより少し時間がかかることがある。

2.2 分析結果出力画面と分析結果の評価

語彙難度分析の結果は図 2 に示すように一覧表とグラフで出力される。画面上部の表 () には語彙レベルごとの頻度および対総語数比率が一覧表示される。この例ではベース辞書に SVL 12000 を指定したため、Level 1 から Level 12 までの結果が表示されている。表中の 01, 02,

Results of Word Level Analysis

Date: 2008.12.25 at 5:9 AM
 Text Title: Sample 1
 Base Dictionary: SVL1200
 Case Sensitive: No
 Count Mode: Token

WL Tag	Word Level	Freq.	%
?	Unknown	50	5.30
01	1,000	766	81.23
02	2,000	47	4.98
03	3,000	35	3.71
04	4,000	19	2.01
05	5,000	8	0.85
06	6,000	4	0.42
07	7,000	2	0.21
08	8,000	6	0.64
09	9,000	3	0.32
12	12,000	3	0.32
-	TOTAL	943	100.00

図2 WLCの分析結果出力画面例

語彙レベルごとの頻度および対総語数比率一覧。ここではSVL 12000を使用したため、Level 1 から Level 12 までの一覧が表示されている。最初の行の Unknown はベース辞書に収録されていない単語を示す。



上記の表を棒グラフにしたもの。Level 1 の単語が圧倒的に多いことが視覚的に確認できる。

Statistics

Total Number of Words	943
Total Number of Word Types	NA
Type/Token Ratio (TTR)	NA
Total Number of Sentences	57
Average Sentence Length (in wrds)	16.3684
Average Word Length (in chars)	4.33762
Average Word Level	1.30435
Var	1.55849
S.D.	1.24839

各種基礎統計量。ここでは語数カウント方法を Token Mode にしたため Word Type 数と TTR は出力されていない。最後の2行は入力テキストの「リーダビリティ」を示す。詳しくは本文解説参照。

Readability Indices

Automated Readability Index (ARI)	7.4
Coleman-Liau Index (CLI)	7.9

[OPTION] To obtain a wordlist of your text, choose one of the following sort types and click submit.

- in order of difficulty level (ascending order)
- in order of appearance in text
- alphabetical order
- frequency order (decending order)

[Toppage](#)

入力テキストの「語彙リスト」を出力したい場合は、ここで左の4つのオプションのいずれかを選択して [SUBMIT] をクリック。語彙リストは MS Excel にインポートしてダウンロードすることができる。

03... という WL Tag (Word Level Tag) はそれぞれ 1000 語レベル、2000 語レベル、3000 語レベルに対応する。指定したベース辞書中に含まれていない単語にはシステム内部で "?" というタグが自動的に付与され、最初の行にある Unknown という欄にその数が返される。表の下にある棒グラフ () は、語彙レベルごとの単語の分布を視覚的に示したものである。

グラフの下部 () には、総語数 (Total Number of Words)、総タイプ数 (Total Number of Word Types)、タイプ・トークン比率 (Type-Token Ratio, TTR)、総センテンス数 (Total Number of Sentences)、平均文長 (Average Sentence Length)、平均語長 (Average Word Length) 等の基礎統計量に加え、平均語彙レベル (Average Word Level) とその分散 (Var. = Variance) および標準偏差 (S.D. = Standard Deviation)、さらに現バージョンで新たに追加された 2 つの「リーダビリティ」指数が示されている。リーダビリティ指数の意味およびその計測方法については後述する。

図2 で例として分析した Sample 1 のテキスト^[5] の場合、総語数の 91.93% (N=867) が基礎4000語以内、2.76% (N=26) が4000語以上、リストに未収録の単語 (Unknown Words) が全体の 5.3% (N=50) となっている。仮に、(かつてそのように想定されていたように) 受験勉強を経た大学新入生の大半が基礎4000語までの語彙を習得済みとすると、このテキストに対する平均的大学生の未習得語率は 8.07% ということになる。ただし、Unknown Words とされた単語の中には否定辞の 'nt ('t) や属格の 's 等の分離処理に伴うエラーやシステムの仕様のために非単語扱いになったもの、および地名・人名等の固有名詞が含まれており、これらの単語を既知語として算入した場合の4000語レンジのカバー率は95.63%、実質未習得語率はおよそ 4.4% ということになる。

Laufer (1989, 1992) は、通常のリーディングにおいて、内容理解に支障をきたさないためにはテキスト中の 95% 以上の語彙が読み手にとって既知語である必要があると述べているが、言い換えれば、未知語率が 5% 以下 (20語に1語) の場合、そのテキストは読み手にとって内容を理解しながら楽しんで読むことができる素材である、ということになる。Krashen (1985) の用語を借りれば、未知語率 5% という数字が、あるテキストが言語習得に必要な “Comprehensible Input” になるかどうかの閾値になると言うことができる。以上のことから、図2 で分析した Sample 1 のテキストは、大学1~2年生を対象にした「英文講読」の教材として、少なくとも語彙の点では “i+1” レベル (学習者の現在の到達レベル +) の適切な難易度にあるものと判断することができる。^[6]

2.3 「語彙リスト」の出力

表2 に示した分析結果出力画面下部には、分析したテキストの「語彙リスト」(Wordlist) を出力するための [OPTION] 欄 () が設けられている。ここでは、次の 4 つのいずれかの方法で HTML 形式の語彙リストを出力することができる。

- in order of difficulty level (ascending order) 難易度レベル順に出力
- in order of appearance in text テキストでの出現順に出力
- alphabetical order アルファベット順に出力
- frequency order (descending order) 頻度順に出力

なお、「語彙リスト」は指定したベース辞書に含まれる単語を対象に、頻度数と難易度タグとともに出力される（WLC ベース辞書については品詞タグも付与される）。指定のベース辞書に含まれていない単語は、語彙リスト末尾にある“Words not in Base Dictionary”という欄に一括して出力される。ちなみに、Sample 1 のテキストでは、以下の単語が SVL 12000 の中に含まれていなかった単語である（括弧の中の数字は頻度数を示す）。

[形容詞 / 副詞] sentient (1) ; traditionally (1) [動詞] dismember (1); minded (1) [普通名詞] conjurer (2); chives (1); immensity (1); powwows (1); nonhumans (1); hummingbird (1); pecan (1); redwood (1); undercurrent (1); wholeness (1) [固有名詞] america (2); denver (1); lesueur (1); meridel (1); mississippi (1); indian (4); oklahoma (3); chickasaw (1); god (4); sunday (1)

これらの単語の中には traditionally や minded, nonhumans などのように語幹からその意味を容易に推測できるものもあるが、通常は sentient, conjurer, chives, immensity のようにごく頻度の低い単語である¹⁷⁾。したがって、当該テキストの学習を始めるに当たって、これらの語彙についてあらかじめ講師が口頭で簡単に解説しておくことで、学生のリーディング活動をより容易にするとともに、語彙習得を促進することが期待できる。もちろん、学生のレベルに応じて、例えば基礎 5000 語以上の単語については一括して語彙リストを作成して学生に配布し、いわゆる pre-reading activity の材料として活用するといった使い方も可能である。本システムでは、こうした使い方を支援するために、HTML ページとして出力される語彙リストの末尾に [Export to Excel] というリンクを設けた。このリンクをクリックすると、当該の語彙リストをエクセルにインポートした上で、ローカルに保存することができる。

3. 「リーダビリティ」の測定

前述のとおり、現在公開されている最新版の WLC には語彙の難易度のほかに、テキスト全体の「リーダビリティ」を測定する機能が追加されている。リーダビリティの自動測定機能を加えるという案は当初からあったが、WLC の最初のバージョンでは同機能の追加は見送ることになった。これは、主として技術的および経済的な問題のためである。

現在、リーダビリティの指標およびその算出方法として、例えば Flesch Readability Score (aka, Flesch Reading Ease Score), Flesch-Kincaid Grade Level, Gunning's Fog Index, SMOG Formula, FORECAST Readability Formula, Powers-Summer-Kearl Formula, Fry Index (aka, Fry Chart) などさまざまなものが提案されているが (Zakaluk and Samuels, 1988; Micro Power & Light Co., N.D.)、いずれも対象テキストの語数、センテンス数、およびシラブル（音節）数を基本にリーダビリティを算出するという方法をとっている。例えば、このうちもっともよく知られ、かつ信頼性も高いとされている Flesch Readability Score は以下の公式で算出される。

$$206.835 - (1.015 \times \alpha) - (84.6 \times \beta)$$

where,

α = average number of words per sentence

β = average number of syllables per word

計算結果は 0～100 までのスケールで示され、60～70 を標準 (Standard) として、得点が高くなるほど「読みやすい」テキストであることを示し、反対に得点が低くなるほど難しいテキストであることを示す。Flesch-Kincaid Grade Level (以下、FKG) は、Flesch Readability Score (以下、Flesch Score) を米国の学年レベルに対応させたもので、次の公式で算出される。

$$(0.39 \times \alpha) + (11.8 \times \beta) - 15.59$$

where,

α = average number of words per sentence

β = average number of syllables per word

なお、FKG の理論的な最低値は -3.4 であるが、これはすべてのセンテンスが単一音節の単語 1 語から構成されているケースを指し、現実にはほぼあり得ない値である。Taylor (2008) によれば、通常のいわゆる “authentic” な (意図的に加工されていない) 英文の場合、FKG の値はほぼ 5.5 から 16.3 の範囲に収まるのが普通であるとされている。

表 1 に、Flesch Readability Score の評価と、各スコアレンジを米国の学年レベルに対応させた一覧表を示す。「学年レベル」とは、それぞれの評価レベルにあるテキストを読むために最低限必要とされる教育レベルを米国式の Grade Level で示したものである。ただし、Grade Level は多少の幅をもって解釈されることが多く、例えば米国の Reference Software International 社がかつて販売していた Grammatik IV というリーダビリティ計測用ソフトウェアでは、FKG = 8 という判定は、ほぼ Grade 6-10 程度の教育レベルに対応する (“requires 6-10 years of schooling”) とされている^[8]。なお、当然のことながら、これは英語の母語話者を対象にしたものである。

Flesch Score	評価	対応する学年レベル(FKG equivalent)
90-100	Very easy	5 年生 (5)
80- 90	Easy	6 年生 (6)
70- 80	Fairly Easy	7 年生 (7)
60- 70	Standard	8-9 年生 (8-9)
50- 60	Fairly Difficult	高校 (10-12)
30- 50	Difficult	大学 (13-16)
0- 30	Very Difficult	大学卒業以上 (17 and above)

表 1 Flesch Readability Score の評価および対応する学年レベル

FKG と同じく、当該のテキストを読むために最低限必要とされる米国の学年レベルとしてリーダビリティを算出する公式のひとつに Gunning’s Fog Index がある。これは、米国の Robert

Gunning が 1952 年に発表したリーダビリティ算出法で、 $0.4 \times (\alpha + \beta)$ という公式で算出される。このうち α は「文当たりの平均語数」を、 β は「3 音節またはそれ以上の音節を持つ単語の総語数に対する比率」をそれぞれ示す。この計算式でも明らかとなり、Fog Index の特徴は音節数が 3 またはそれ以上の単語を “big word” とし、その比率をリーダビリティ判定の重要な要素とした点にある。なお、Fog Index は FKG よりいくらか高めになることが知られている（染谷 1986/1994, Vol. 3, p. 150）。

本稿では、こうしたさまざまな指標についてひとつひとつ詳しく取り上げることはしないが、既存のリーダビリティ公式を Word Level Checker のようなオンライン上のシステムで採用する際の一番の問題点は、いかにして「音節数」を正確にカウントするかという点にある。コンピュータに入力データの音節数を正しくカウントさせるためには、まずはすべての英単語の（あるいは一般化規則に当てはまらない例外的なものについて網羅的にその）音節数を手動でカウントし、これを一覧表にしたデータベースを作成する必要がある。単純に母音の数をカウントするという方法では、仮にいくつかの例外規則を盛り込んだとしても、正確な測定は望めない^[9]。このようなリストの作成は時間と労力さえかければ可能な作業ではあるが、一教員が授業の合間を縫ってコツコツ作成するというレベルの作業ではない。外注すればいいのだが、その予算はない ということで、結局、2006 年の段階では、リーダビリティの自動測定機能の追加は見送らざるを得ないということになった。

3.1 「音節数」に依拠しないリーダビリティの計測法：ARI と CLI

音節数のカウントということがネックになっているとすれば、音節数に基づかないリーダビリティの計測方法を採用すればよいということになる。実は、その後の調査で、そのような方法が少なくとも 2 つ、すでに提案されていることがわかった。ARI (Automated Readability Index) および CLI (Coleman-Liau Index) と呼ばれるものである。

ARI は、もともと米国防軍内での文書（とくに技術文書やマニュアル）の作成・評価を支援する目的で 1960 年代後半に提案されたもので、最初からコンピュータでの自動処理を念頭に入れて考え出されたものである。ARI の算出式は以下のとおりである (Smith and Senter, 1967)。

$$\text{ARI} = (4.71 \times \text{CPW}) + (0.5 \times \text{WPS}) - 21.43$$

where,

CPW = average number of characters per word (= total number of C divided by that of W)

WPS = average number of words per sentence (= Total number of W divided by that of S)

CLI は、Meri Coleman と T. L. Liau が 1975 年に考案したリーダビリティ測定法で、上記の ARI と同じく、コンピュータでの自動処理を前提に、他の大多数のリーダビリティ・フォーミュラと異なり、「音節数」に拠らずに対象テキストのリーダビリティを測定するものである (Coleman and Liau, 1975)。CLI の算出式は以下のとおりである。^[10]

$$CLI = (5.89 \times CPW) - (0.3 \times (100 \div WPS)) - 15.8$$

where,

CPW = average number of characters per word (= total number of C divided by that of W)

WPS = average number of words per sentence (=Total number of W divided by that of S)

ARI と CLI の計算結果は、いずれも FKG と同じく米国の学年レベルとして算出される。Taylor (2008) によれば、ARI は FKG および CLI よりやや高めに算出される傾向があり、また、技術的な文書を対象にした場合、CLI は一般に FKG や ARI よりも低くなるとされている。

以上、音節数のカウントという手順を経ずにリーダビリティを測定する方法があることを確認した。ARI と CLI はいずれの場合もテキストデータの形態的な指標、つまり（スペースで区切られた）単語の数とこれを構成する文字の数、および（特定の記号によって明示的に区切られた）文の数のみによってリーダビリティを測定する方法であり、コンピュータによる自動処理にもっとも適した手法であると考えられる。問題は、これらの方法がどの程度信頼性があるかということになる。

3.2 FKG, ARI, CLI の相関関係

本節では、FKG を基準に、ARI および CLI との相関関係について検討する。前述のとおり、FKG（およびその前提となる Flesch Score）はリーダビリティ評価のスタンダードとして定評があるものであり、これとの相関関係が有意に高ければ、ARI および CLI の信頼性も高いことが証明されることになる。調査に当たって、FKG の算出は前述の Grammatic IV という市販のソフトウェアに拠り^[11]、ARI および CLI の算出については自作のコンピュータプログラムを使った。なお、最新版の WLC に実装されているプログラムは Java 言語で記述されているが、これに先立って、簡易言語 AWK で実験用プログラムを作成した。プログラムのアルゴリズムは全く同じのものであり、どちらを使っても得られる結果は同じである。

文の定義と文の数のカウント

前記の実験用プログラムでは、「文」の数をできるだけ正確かつ統一的な方法でカウントするために、以下のようなコントロールをプログラム中に組み込んだ。

- 「文」の定義 = 基本的には . ! ? のいずれかで終わっている文字列を「文」とする。
- 箇条書きが含まれている場合、行頭の「英数字 + ペリオド」(/^[A-Za-z0-9]¥. / という正規表現で表される文字列) を文末記号として認識しないように回避。
- 文中・文尾のエリプシス記号 (3点ドット) を文末記号として認識しないように回避。
- 文末で2重に使われた感嘆符と疑問符は1回のみ文末記号として認識。
- 間にペリオドを挟む2~5文字の略語、およびペリオドのあとに(不正に)スペースの入った2~5文字までの略語は、適切に「略語」として認識されるようにする(略語ペリオドを文末記号と誤認識するのを回避)。
- 引用符で囲まれた文字列が前記の文末記号のいずれかで終わっている場合、地の文と分けてカウントする。

- コロンの後が「スペース+大文字」で始まっている場合、コロンを文末記号とする（小文字の場合は文を分けない）
- メールアドレスと URL は便宜上それぞれ 4 文字の 1 語扱い(MAIL と HTTP) として処理。

なお、上述のとおり、ピリオドで区切られた略語が文中に置かれているケースについては略語ピリオドが文末記号と誤認識されることを回避するようにしたが、文末位置に略語が置かれているケース 例えば The organization is known as the [N.S.A. This] is... のような文では、[] の部分が His name is [D.H. Smith] and he is... のように文中の略語要素として使われているもの（「大文字+ピリオド+スペース+大文字」という連続の文字列）と形態上の区別が付けられない。同じく、文中の Mr., Mrs., Ref., Prof., Assoc. といった形式の略語は通常文の文末と同じ形式になることがあるため、単純なパターンマッチングによる処理ができない。前者の場合は適当な回避方法がないが^[12]、後者についてはこのようなタイプの略語を網羅的にリストアップし、ひとつずつつぶしていく方法を採用した。現在のプログラムにはこの形式の主な略語が 124 例、登録されている。完全とはいえないまでも、実用的には必要十分な数であると考えられる。

テスト用の英文サンプルについて

テスト用の英文サンプルには、(1) 筆者の勤務校で実際に使用している英文講読用教材から抜粋した英文エッセイ 4 本、(2) リーダビリティ評価用に筆者が授業で使用しているビジネスリーダーの例文 7 本、さらに (3) リファレンス用のテキストとして米国大統領の就任演説を 2 つ使用した。米国大統領の就任演説は、部分的にやや難かしい箇所が出てくるものの、もともと一般国民向けに作成されたもので、英語国民にとってごく標準的な（かつ、一定の格調を確保した）英文であることから、日本の大学生向けの英文講読用テキストの“Fitness”を測る対照テキストとして格好の素材であると判断した。なお、(1) および (2) については、主観的評価による難易度判定を、「易しめ (easy)」「普通 (average)」「やや難 (difficult)」「難 (very difficult)」の 4 段階にわたって加えた^[13]。(3) はこのうちの「普通 (average)」という判断であるが、これは日本の大学生にとっての「到達目標レベル (target level)」と考えてもよい。

テストサンプルの分析結果

表 2 に、各サンプルのリーダビリティ評価の結果を示す。なお、注 11 で述べたとおり、FKG のスコアについては少数点以下が省略されている。

この結果を見るかぎり、FKG, ARI, CLI のスコアはほぼ一定の傾向を示しているように思われる。すなわち、「易しめ (easy)」と判断されたテキストのスコアは 4.3 から 8.8 (STD = 0.4 ~ 0.9) と比較的 low、普通 (average) が 8.0 から 11.9 (STD = 0.10 ~ 1.16)、「やや難 (difficult)」および「難 (very difficult)」と判断された 4 つのテキストは、いずれも 12.1 から 15.7 (STD = 0.23 ~ 0.95) と高い数値を示していることがわかる。各スコアの標準偏差の平均は 0.56 (STD = 0.35) で、数値が比較的まとまっていることを示している。これは、図 3 のレーダーチャートを見ても明らかである。

N	TEXT ID.	リーダビリティ指数					TEXT ID.	テキストタイプと主観的評価による 難易度 (RF = Reference Text)
		FKG	ARI	CLI	Ave.	STD		
1	Sample 1	8	7.4	7.9	7.8	0.32	エッセイ：講読用教材からの抜粋（易）	
2	Sample 2	14	13.6	13.6	13.7	0.23	エッセイ：同上（やや難）	
3	Sample 3	10	10.8	9.6	9.6	0.61	エッセイ：同上（普通）	
4	Sample 4	14	15.0	14.7	14.6	0.51	エッセイ：同上（難）	
5	Speech 1	11	11.1	11.2	11.1	0.10	米国大統領就任演説（Bush 2006）(RF)	
6	Speech 2	10	10.3	10.2	10.2	0.15	米国大統領就任演説（Clinton 1997）(RF)	
7	BL-07	10	9.3	9.9	9.7	0.38	ビジネスレター（普通）	
8	BL-08	8	7.0	8.8	7.9	0.90	ビジネスレター（易）	
9	BL-13	10	9.2	8.0	9.1	1.01	ビジネスレター（普通）	
10	BL-19	11	9.6	11.9	10.8	1.16	ビジネスレター（普通）	
11	BL-23	14	15.7	15.6	15.1	0.95	ビジネスレター（やや難）	
12	BL-30	13	13.1	12.1	12.7	0.55	ビジネスレター（やや難）	
13	BL-60	5	4.3	5.0	4.8	0.40	ビジネスレター（易）	

表2 サンプルテキストのリーダビリティ指数対照表

* STD Ave. 0.56 (STD = 0.35)

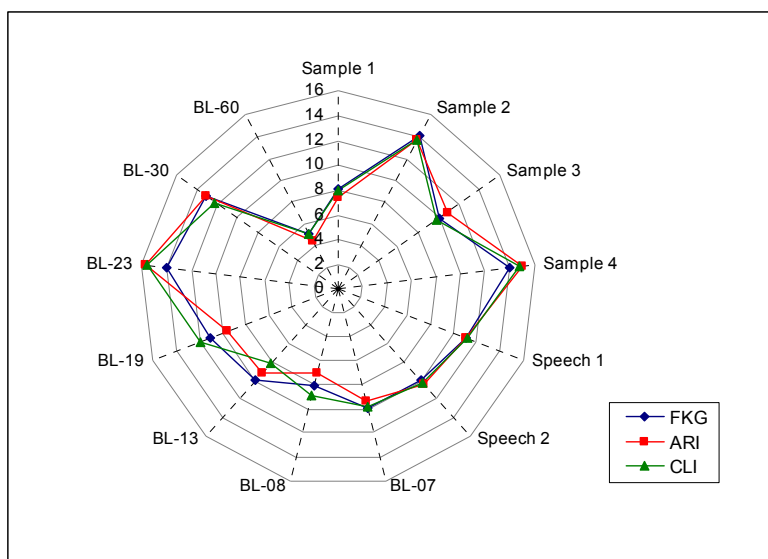


図3 リーダビリティ評価のレーダチャート

なお、一般に、テキスト量が少ないものについては評価がやや不安定になる傾向があるが、本調査でもビジネスレターの例文（とくに BL-23, BL-19, BL-13, BL-08）についてその傾向が見られる。したがって、実際のテキストの分析に当たっては、できるだけ 1000 語程度またはそれ以上のまとまった量を対象にするのが望ましい。ちなみに、Speech 1 は総語数 6803 語、Speech 2 は 5283 語、Sample 1 ~ Sample 4 についてはそれぞれ 933 語、932 語、1128 語、513 語となっている。

表3に、FKG, ARI, CLIの3つのリーダビリティ指標間の相関係数行列を示す。一般に、相関係数 r は $0.8 < |r|$ のときに「強い相関あり」と判断されるが、表3を見ると、FKG-ARI間の相関係数が0.9775、FKG-CLI間の相関係数が0.9539、ARI-CLI間の相関係数が0.9484と、いずれも強度の相関関係を示していることがわかる。

(r)	FKG	ARI	CLI
FKG	1	0.9775	0.9539
ARI	0.9775	1	0.9484
CLI	0.9539	0.9484	1

表3 リーダビリティ指標間の相関係数 (r)

図4は、3つの指標によるリーダビリティ指数 (Grade Level) の散布図に、回帰直線 (線形近似) を当てはめたものである。これを見ると、FKGを基準とした場合、ARI (点線) と CLI (一点破線) はいずれも、われわれが「到達目標レベル」とした Grade Level 10~11 あたりを境界に、スコアの差がより強調されて (低いものはより低く、高いものはより高く) 算出される傾向があることが伺える。また、ARIはもともと技術文書のリーダビリティを計測することを目的としているため、子供向けのごくやさしい文章や短いセンテンスからなる会話テキスト (例えばFKGで2前後またはそれ以下のもの) を対象にした場合、不当に低い評価が出る傾向がある。スコアの評価に当たっては、前述のテキスト量の問題とともに、この点も考慮に入れるべきであろう。

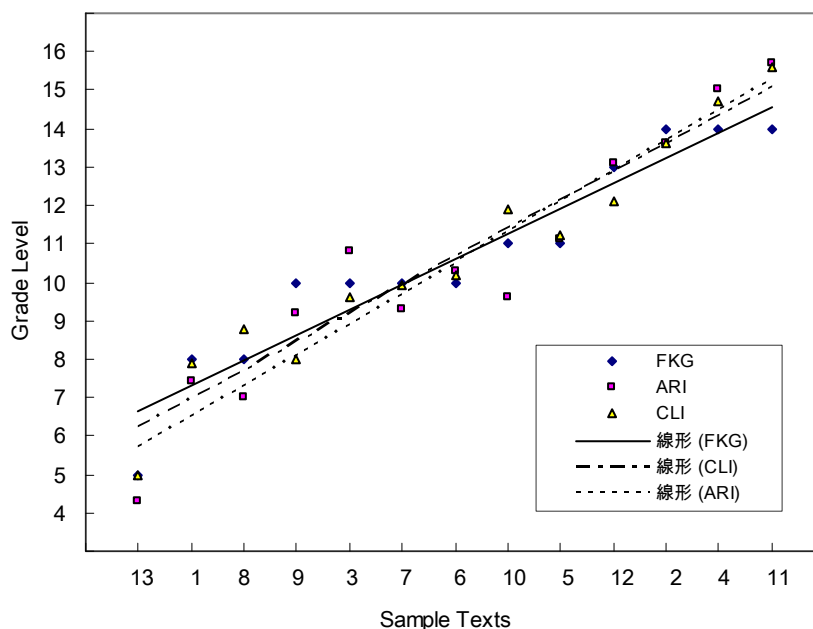


図4 リーダビリティ指数の散布図と回帰直線

4. WLCの応用可能性と今後の課題

前節では、FKG, ARI, CLIの3つのリーダビリティ指標間に強度の相関関係があることを確

認した。結論としては、WLC に新たに追加された ARI と CLI はリーダビリティ評価の指標として十分に実用に耐えるものであるということになる。ちなみに、日本の大学生用の講読テキストとして使用する場合の ARI/ CLI スコアの評価の目安は表 4 のとおりである。

ARI/CLI スコア	評価の目安
5 or less	Too Easy
7 - 6	Fairly Easy
8 - 9	Average
10 - 11	Fairly Difficult
12 - 13	Challenging
14 or more	Too Difficult

表 4 ARI/CLI スコアの評価の目安

日本の大学生用の講読テキストとして
使用する場合の目安

ただし、当然のことながら、これらの指標によるリーダビリティ評価は（他の指標と同じように）単にテキストの表面的・形態的な特徴のみによってリーダビリティを判断しているため、たとえば A green boy is a dog and cat in the house of fire. のような全く意味を成さない文字列や Four king sons had the. のような非文でも、「読みやすい文」(!) という評価になる。反対に、長めの地名や人名が多く出てくる文章や、カンマやセミコロン等で付加的な語句を重層的につなげたものは、仮に文章としてはごく簡単に理解できるものであっても、リーダビリティ評価は低くなる。したがって、機械的な適用は避け、あくまでも参考データとして利用すべきものである^[14]。ただし、本稿で紹介したような語彙難易度評価やリーダビリティ評価は、その限界を自覚しつつ適切に利用すれば、教育研究のさまざまな側面で大いに役に立てることができるものである。以下、本稿のまとめとして、WLC の教育研究分野への応用可能性について若干の提言を行い、最後に現在の WLC の問題点と今後の課題について述べる。

4.1 WLC の教育研究分野への応用について

WLC の応用分野として、まず第 1 に挙げられるのは、そもそもの開発の契機となった「英文講読」の授業への応用であろう。このほかに、「英文ライティング」分野への応用、および入試問題の作成と品質管理への応用が考えられる。

4.1.1 「英文講読」授業への応用

いわゆる英語科目としての「英文講読」の授業は、通常、どの大学でも非常勤講師が担当することが多い。教材の選定や指導内容についても、大学側でごく大雑把な方針を示すことはあっても、一般には「教育実践不介入」の原則から、具体的なことについてはすべて講師任せにしているのが普通である。もちろん、すべてうまくいっている場合はこれでいいわけだが、実

際には、学生のレベルや授業の設置意図を無視した恣意的な授業や、講師の趣味に偏りすぎた内容の教材を使っている例がないわけではない。毎年の学生による授業評価にもそのようなコメントが散見する。

WLC は、そのような背景から、各教員による自己チェックのための支援ツールのひとつとして開発されたものである。もちろん、この試みは非常勤の先生方の教育実践に介入しようというものではなく、あくまでも各教員が自らの授業で使う教材の語彙的な難易度やリーダビリティを客観的に測定し、その結果をよりよい授業の運営に役立ててもらおうというものに過ぎない。したがって、筆者の勤務校においてもその使用・不使用はまったく自由であり、何ら拘束力を持つものではない。この点は筆者の勤務校で毎年行われている講師打ち合わせ会においても繰り返し強調されていることである。なお、この「打ち合わせ会」では、WLC の利用法について詳しく解説するとともに、講師の方から提供される教材サンプルの分析例（「テキストプロファイリング・レポート」）を作成・配布しているが、参考までにその抜粋を巻末資料に示す。

この巻末資料に挙げた例は、表 2 の Sample 4 について、より詳細に分析したものであるが、テキストの“Fitness”を測定するための「自己診断ツール」として WLC がどのように役に立つをよく示しているものと思われる。なお、このレポートを作成した 2007 年度の時点ではリーダビリティ評価は含まれていないが、ここでは参考までに ARI と CLI による評価、および参考値として FKG のスコアを追加した。

このレポートにあるとおり、Sample 4 のテキストは語彙の難易度（4000 語レンジの語彙のカバー率 = 82.59%、同推定未習語率 = 17.41%）、およびリーダビリティ評価（FKG = 14; ARI = 15.0; CLI = 14.7）のいずれの点でも突出した難易度のテキストであり、日本の平均的な大学生を対象にした通常の「英文講読」用の教材としては難易度が高すぎると考えられる。したがって、分析者の提言としては、「同一の内容を扱った、より難易度の低いテキスト（Grade Level 評価で最大 10~11 程度のもの）を選択することが望ましい」としている。もちろん、その上で、あくまでもこの教材を使用するか、あるいは同一内容のより難易度の低いものに換えるかは担当講師の判断に任せることになる。

もちろん、難易度が高くても、教育的な観点からその教材を使うのが最善であるということも十分にあり得る。ただし、その場合には、本稿の第 2 章 3 節で述べたとおり、当該テキストの学習を始めるに当たって、あるいはこれに並行して、一定のレベル以上の難易度を持つ語句について講師が口頭で簡単に解説したり、「語彙リスト」を作成して学生に配布し、pre-reading ないし post-reading activity の材料として活用するなどの工夫が必要になるだろう。すでに述べたとおり、WLC にはこうした使い方を支援するための仕組みが用意されている。

4.1.2 「英文ライティング」分野への応用

2 つ目の応用分野として、「英文ライティング」分野への応用が考えられる。通例、学生の書く英文（エッセイやジャーナルなど、一定の量があるもの）は、自分がすでに知っていて習熟度の高い語彙の範囲内で書くため、語彙的なバラエティに欠けることが多いが、WLC を使って学生の書いた英文データを分析し、併せて難易度付きの語彙リストを作成することで、その

学習者の L2 レキシコンのおそよの幅と多様性、あるいは語彙使用における傾向といったものを把握することができる^[15]。こうして得られたデータをもとに、例えば同じ単語を繰り返し使っていないか、別のより適切な表現に言い換えることが可能か 可能なら、どういう選択肢があるか等々といった点について、具体的かつ個別的な指導をすることで、よりいっそうの教育効果が期待できると考えられる。また、学期開始時点と学期終了時の語彙使用を比較することで、学生個人およびクラス全体の長期に亘る変化の有無を観察することもできる。こうして得られたデータや知見は、教師にとって貴重な教育的データとなることは間違いのないところであろう。

もちろん、語彙だけでなく、学生の書いた英文のリーダビリティを測定し、これをライティングの指導に役立てることもできる。例えば、文章をただらと続けていないか、こま切れ的なセンテンスになっていないかといった点に学生の意識を向けさせるだけでも、文章の質を大きく向上させることができる^[16]。その上で、文章の内容的なことや論理性、結束性といった側面については、これまでどおり教師が個別に指導を行っていけばよい。

ただし、こうしたことを円滑かつ効率的に行うためには、学生が提出する英作文データをすべて電子データとして収集・保存しておく必要がある。もっとも手軽な方法としては、エッセイなどの提出はすべて電子メールで行い、ジャーナルの場合はインターネット上に掲示板（BBS）やブログを立ち上げ、すべてこれに書き込ませるようにすることで、とりあえず、すべてのデータを電子データとして取り出すことが可能になる。個々の教員が自分の担当するクラスのみを対象にする場合はこれで十分であるが、ひとつの学科や学部、あるいは全学を対象にする場合はこのような原始的な方法の限界は明らかであり、本来は、学生の作成した英文データを統一的な方法で収集・保存し、これを「コーパス」として利用できるようなシステムを用意しておくべきであろう。筆者の勤務校においては今のところそのようなシステムがないため、毎年の授業で発生する貴重なデータが使い捨て状態になっているが、よりよい教育実践を支援するために、いずれそのようなシステムを立ち上げるべきであると考えられる。この点で、早稲田大学国際教養学部の "The SILS Learner Corpus of English" と呼ばれる教育研究実践例 (Muehleisen, 2007) は大いに参考になるところである。

4.1.3 入試問題の作成と品質管理への応用

WLC のもうひとつの応用分野として筆者がとくに興味を持っているものに、大学入試問題の作成およびその品質管理への応用がある。大学が独自の入試問題を作成する場合、通常は当該学部・学科の教員が交代でその任に当たっているため、出題形式は一定でも問題の難易度が年度によってかなり異なってくることもある。また、入試問題にしばしば高校生の力をはるかに超えた難問が出されることがあるのは周知の事実である。

中條・長谷川 (2004) は 2002 年度に実施された大学入試のうち、筆者の勤務校である青山学院大学を含む合計 26 大学 40 学部の入試問題における英文読解問題の難易度を分析している。分析のパラメタは、入試問題に対する「中高英語教科書語彙」のカバー率と、入試問題のリーダビリティ (Grade Level による評価) の 2 点である。 については、現行の中学校および

高等学校教科書に使用されている語彙を網羅的に収集したコーパスを作成した。高校用の教科書については高等学校修了時の語彙レベルの上限に相当する例として、難易度の高い教科書シリーズを採用している。作成された「中高英語教科書コーパス」の異語数は 3,098 語（延べ語数は 43,772 語）であった。JACET 8000 との相関については触れられていないが、JACET 8000 のうちの 3000 語レベルまでの基礎語彙が収録されているものと考えてよいであろう^[17]。リーダビリティについては本稿第 3 章冒頭で言及した各種リーダビリティ指標を含む 9 つの指標のうち、大学入試問題のリーダビリティ判定にもっとも適していると判断された Flesch-Kincaid Grade Level, FORCAST Formula, Fry Index の 3 指標の平均を用いている。

中條・長谷川（前掲書）ではその結果を一覧表にして提示しているが、青山学院大学（学部学科名は不詳）については、当該入試問題に対する中高教科書語彙のカバー率が 88%（全体平均 90.9%）、リーダビリティは 13.1（全体平均 9.5）となっている。ちなみに、同年のセンター入試の英文読解問題の中高教科書語彙カバー率は 94.2%（1993-2002 年の平均は 94.7%）、リーダビリティは 8.4（同平均 6.2）である。

リーダビリティが 13.1 というのは全 40 学部のうち上位 3 番目の難易度である^[18]。もちろん大学の「格」と入試問題の難易度とは何の関係もない。これに対して、調査対象となった 40 例の約半数に相当する 23 学部がリーダビリティで 8~10 あたりに集中している。これは、高校生を対象とした入試問題としてごく適切な判断というべきであろう。リーダビリティ 13 というのは、大学入学時ではなく、卒業時に（かつ、一部の優秀な学生において）到達していることが期待されるレベルである。中條・長谷川もその結論で述べているとおり、大学入試問題はその社会的な影響の大きさを十分に考慮に入れた上で、「受験者の習熟度にふさわしいレベルの英文」を出題すべきであり、出題者としてはその点に十分に配慮して作問に当たるべきであろう。

もちろん、入試問題にはそれぞれの作問者の出題意図があり、統計的な指標のみでその是非を論ずることはできないが、受験者の習熟度に応じたレベルの英文においてその出題意図が表現されているほうが好ましいことは言うまでもない。入試問題を作成するに当たって、一種のスクリーニング装置として WLC を活用することで、問題の難易度が必要以上に高くなりすぎないように調整することができる。例えば、JACET 8000 のうち、4000 語レベルを超えるもの（ただし、これはそれぞれの大学の求めるレベル規定による）については問題末尾に語注を加えたり、設問とは直接関係のない箇所については一定のリーダビリティ内（例えば 10 から 11）に収まるように文章を調整するなどの客観的な基準を設けた上で問題の調整に当たることによって、年度ごとのばらつきや出題者ごとの判断の揺れを最小限に抑えることができる。その上で、難易度の高い文や語彙を意図的に一部残しておくことは「試験」という性格上、当然、許容されてしかるべきものである。

4.2 WLC の問題点および今後の課題

本稿では、WLC の機能について簡単に解説した上で、現行バージョンに新たに追加された ARI および CLI がいずれもリーダビリティ指標として一定の制限内で十分に実用に耐えるものであることを検証し、併せてその応用可能性について議論した。

一定の制限というのは、リーダビリティ評価そのものが対象テキストの表層的な特徴のみによって文章の読みやすさを判断するものであるという限界に加え、現在の WLC にはある条件下において不適切な語分割や文分割をする可能性があるというパーシング上の問題が残っていることを指す。このうちパーシング上の問題については今後改良の余地があるが、現時点では分析サンプルの語数を増やすことで問題の影響を最小限に抑えることができる。本文中でも述べたとおり、分析サンプルのサイズは大きいほうが安定した結果を得られるため、できるだけ 1000 語以上のサンプルを使うことを勧めたい。

なお、今後の課題として、ベース辞書への品詞タグ情報付与をできるだけ早い時期に完了させたいと考えている。現時点では JACET 8000 と WLC ベース辞書（見出し語数約 3 万 5000 語）については収録されているすべての単語について作業が完了しているが、SVL 12000 への品詞タグ付けが未完了になっている。この作業が終われば、単なる語彙リストのほかに、分析対象テキストのそれぞれについて品詞ごとの頻度情報やコリゲーション情報（colligation = 品詞間の共起・連鎖関係）を必要に応じて出力することが可能になる。また、リーダビリティ評価についても、既存のほとんどのリーダビリティ指標が音節数を主要なパラメタのひとつとして取り上げていることから、音節数をコンピュータで正確に自動カウントするためのアルゴリズムについても検討を重ね、将来的には ARI および CLI 以外の指標も加えて複数の指標間の比較ができるようにしたい。これが 2 つ目の課題である。次回のバージョンアップの際には、是非ともこれらの課題が実現できる状況になっていることを期待したい。

巻末資料 英文購読用テキストのプロファイリングレポート（抜粋）

英文購読用テキストの難易度評価について（2007 年度 講師打ち合わせ会参考資料）

[テキスト本文] (Sample 4)

Although Socrates is principally associated with the spoken word, he was born to the visual arts. His father, Sophronistlos, was a stone-worker (lithourgos, marmorarius), and he himself is reported to have worked as a sculptor early in life. Ancient literary sources report that he carved a group of the Chnrites (Graces) that stood at the entrance to the Athenian alvopolis as well as an image of Hermes Propylaios. Attribution of actual surviving artworks to Socrates -- just like words -- remains highly problematic, of course, but multiple marble versions of both compositions are preserved. Their popularity in antiquity is likely the result of (mis)attribution to the philosopher, for Socrates was not a rare name; and homonymous craftsmen, among numerous others, are mentioned in ancient texts. . . . (以下、省略)

[語彙の難度分析の結果]（要約のみ。表・グラフは省略）

- 4000 語レンジ（既習レベル）の語彙のカバー率 = 82.59%
- 同推定未習語率 = 17.41%（サンプル中の 529 語のうち約 92 語 各センテンスごとに平均 4 語 が未習得語彙と想定される）

SVL 12000 に未収録の単語 (unknown words) から固有名詞と変換エラーの 17 語（3.2%）を既知語として参入した場合のカバー率。

[予想難語リスト]

（省略）

[リーダビリティ評価]

Total number of words = 513; Total number of sentences = 22;

Ave. sentence length = 23.3182 words; Ave. word length = 5.4152 characters;

Flesch-Kincaid Grade Level = 14 (Flesch Reading Ease Score = 34) ;

ARI (Automated Readability Index) = 15.0; CLI (Coleman-Liau Index) = 14.7; Ave. Grade Level = 14.57

Remarks: This text is difficult for most readers. It requires at least 14 years of formal education to understand. (This assessment assumes that the target reader is a *native speaker* of English).

[コメント] このテキストは今回分析したサンプルの中でもっとも難易度の高いものである。修正後の実質未習得語彙比率 (17.41%)、文当たりの平均語数 (23.3 words)、長文比率 (35.4%)、受動態比率 (50%)、Type-Token 比率 (52%) など、どのパラメタを見てもこのテキストの難易度は突出したものであり、母語話者にとっても読解には約 14 年の正規教育（大学 2 年程度の教育レベルに相当）を必要とするテキストとなっている。したがって、日本の大学生を対象にした通常の「英文購読」用の教材としては難易度が高すぎると考えられる。同一の内容を扱った、より難易度の低いテキスト（Grade Level 評価で最大 10～11 程度のもの）を選択することが望ましい。ただし、専門ゼミのテキストとして、教師の適切な指導のもとに使用する場合は必ずしもこの限りでない。

【注】

- [1] 染谷 (1998) 参照。同論文は次のサイトで閲覧可能 ([Online] http://www.someya-net.com/kamakuranet/wlc/wlc_manual.html)。なお、Java CGI への移植に当たっては、青山学院大学文学部英米文学科の卒業生で、現在、ソフトウェア開発を中心とする IT 企業を経営しておられる米村歩氏の助力を得た。
- [2] 本論文は、青山学院大学の在外研究員制度の適用を受けて、筆者が 2008 年 4 月 1 日から 1 年間、オークランド大学 (ニュージーランド) に滞在中に執筆したものである。ここに記して関係者への感謝の言葉としたい。
- [3] ちなみに、大学英語教育学会からは JACET8000 の使用に当たって 30 万円の使用料を請求された。もともと「広く教育研究用に資する」ことを目的とし、かつすでにそのような趣旨で公開されているデータを元に作成されたこの種の学術的な成果物に対し、一律に「使用料」を請求するという判断は、同じ大学人としていささか納得しがたいところである。早急に見直しの決定がされることを期待したい。(株)アルクからは無償の使用許諾をいただいた。ここに記して、改めて感謝の意を表したい。
- [4] 基底語化 (lemmatization) とは、各単語の変化形や派生形を元の形 (base form = lemma) に戻すことを指し、本プログラムでは染谷 (1998) で作成したレンマリストを使って基底語化を行っている。なお、このレンマリストは特定ジャンルの語彙を処理するために作成されたもので、必ずしもすべての単語とその変化形が収録されているわけではない。したがって、入力テキストによっては、一部、基底語化されない単語や不適切な処理をするケースがある。
- [5] 青山学院大学文学部で 1~2 年生向けの「英文講読」の授業で使われている教材からの抜粋で、担当講師による難易度の主観的評価は「易しめ (easy)」となっている (表 2 参照)。
- [6] もちろん、内容の理解には語彙以外のさまざまな要素が関係してくるが、その中でも語彙の難易度がもっとも大きな要因になっていることは疑問の余地のないところである。ちなみに、清川 (1996) は、読解力テストの誤答率と単語の難易度の相関係数はおよそ 0.6~0.7 であると報告としている。なお、ここでは対象とする学生の大半が基礎 4000 語 (高校までの学習語彙約 3000 語 + 1000 語) 程度の語彙を習得済みという前提で話を進めているが、この前提を最近の大学生 (といっても、すでにそのような状態が 10 年ほども続いているように思われるが) について、そのまま当てはめてよいかは大いに疑問である。酒井 (2006) は、一部の大学では「英検 3 級や 4 級程度の大学生」(つまり中学卒業程度あるいはそれ以下の英語力しかない大学生) もかなり存在し、授業の運営に大きな支障をきたすほど基礎力が低下してきていることを報告しているが、いわゆる偏差値の高い有名大学の英語専攻の学生でさえ、その全体的な語彙力・文法力の低下傾向は顕著であり (染谷 2007)、基礎 4000 語を習得済みという前提は、もはや無条件の前提とはできない状況になっているものと思われる。
- [7] なお、固有名詞は本来、America, Denver, Lesueur (または Le Sueur = 人名) のように語頭を大文字で表記すべきであるが、本稿で取り上げた例 (図 1 参照) ではケース指定を "No" のまま実行しているため、すべての単語が小文字化されて処理されている。したがって、ここではベース辞書に大文字で登録されている単語 (America, Sunday, God など) がすべて Unknown Words としてカウントされている。
- [8] Flesch Reading Ease Score, Flesch-Kincaid Grade Level および Gunning's Fog Index の解説と計算式は染谷 (1986/1994, Vol. 3, pp. 148-153) から引用した。表 1 も同書から一部変更して転載したものである。なお、Flesch Reading Ease Score と Flesch-Kincaid Grade Level (FGK) は MS Word の文章校正機能の一部

として組み込まれており、以下の手順で使用することができる。

1. MS Word のメニューから [ツール]→[オプション] を選択。
2. [スペルチェックと文章校正] タブをクリックし、「文章校正」のところにある「文章の読みやすさを評価する」にチェックマークを入れる。
3. MS Word に分析対象の文章を貼り付け、メニューから [ツール]→[文章校正] を選択。
4. 「読みやすさの評価」結果が別ウィンドウに表示される。

結果表示ウィンドウには、Flesch Score と FKG とともに、統計量として単語数、文字数、段落数、文の数、文当たりの語数、単語当たりの文字数等が表示される。ただし、「音節数」は表示されないため、評価結果を検証することができない（同じテキストを異なったソフトウェアで評価した場合、同一のリーダビリティ指標を使っても必ずしも同じ結果にならないことがあるが、これは主として音節数のカウント方法とセンテンス分けのためのアルゴリズムが異なるためであると思われる）。

このほか、Flesch Score と FKG の算出に特化したオープンソースのフリーウェアとして **Flesh 2.0** があり、次のサイトからダウンロードすることができる（→ <http://flesh.sourceforge.net/>）。こちらは語数、文数のほかに音節数も表示されるため、評価結果を検証することができるようになっている。なお、Flesh 2.0 は Java ベースのプログラムであるため、使用に当たっては別途、次のサイトから Java をインストールする必要がある（→<http://www.java.com/en/download/manual.jsp>）。

- [9] 母音の数を基準にした自動処理でも 95% 前後の精度は確保できると考えられるが、後述のように文のカウントについてもある程度のエラーを見込む必要があることから、このような方法をとった場合、全体の信頼性が著しく下がってしまう可能性がある。（いくつかのカウント規則を組み込んだ上で例外的な単語を別途リストアップしておけば、十分な精度を持った自動化プログラムが可能だと思われるが、これは今後の課題としたい。）
- [10] CLI の算出式にはいくつかのパリエーションがある。この式は Wikipedia (http://en.wikipedia.org/wiki/Coleman-Liau_Index) および Readability.Info (<http://www.readability.info/info.shtml>) に掲載されている記述によったもので、いわゆる「簡略版 (simplified version)」とされているものである（ただし、簡略版にも別式がある）。なお、本稿で取り上げた各種公式に用いられている係数の算出プロセスについては現在のところ確実な資料が得られていないが、通例、係数の決定は、ある一定の基準に従って収集されたデータを対象に、リーダビリティに影響すると考えられる各種のテキスト要因（例えば 1 文の平均語数や音節数、文字数など）を独立変数とし、学年レベルを従属変数として重回帰分析を行い、有意であるとして残った変数の係数から最終的な計算式を作成する、という手順を踏むものと考えられる。なお、重回帰分析は複数の独立変数（この場合は各種テキスト要因）からひとつの従属変数（この場合は学年レベル）を予測するための式を作る際に用いる統計手法で、独立変数が未確定の場合はステップワイズ法と呼ばれる手法を使う。
- [11] Grammatic IV（および類似のソフトウェア）による評価はいずれも小数点以下を切り上げた整数で示されるため、他の指標との正確な比較ができにくい。また、シラブルのカウント方法を含め、計算の内部プロセスを検証することができない。出来合いのソフトウェアを使う際の大きな問題点のひとつである。比較に当たっては、本来、小数点以下の桁数を揃えるべきであるが、以上のような事情で本稿では FKG のスコアについては小数点以下を割愛した数字をそのまま使うことにする。

- [12] ただし、あらかじめ入力テキストを編集する（例えば The organization is known as the NSA. This is... のように、文末記号以外のピリオドを削除する）ことで問題を回避することができる。
- [13] (1) はデータを提出していただいた講師の方による主観的評価、(2) は本稿の筆者による判断である。なお、(2) のビジネスレターの例文のリーダビリティについては、染谷（1986/1994, Vol. 3, Figure 95, p. 151）で詳しく議論している。
- [14] 現在のリーダビリティ評価では、対象テキストの内容的な複雑さ（または簡明さ）や論理性、文法性、使用語彙の具体性・抽象性、結束性や一貫性、あるいはディスコースの文化的・政治的・思想的な偏りや特徴、さらに読み手の読解意図や動機といった、本来、テキスト読解に大きな影響を与える質的要素は一切考慮されておらず、機械的に適用すると大きな間違いを犯すことになる。こうしたことは、リーダビリティ評価を採用するに当たって当然の前提ではあるが、ここで改めて確認しておきたい。
- [15] 語彙の多様性を測るひとつの指標としてタイプ・トークン比率 (Type-Token Ratio) がある。タイプ・トークン比率は、一般には 1000 語単位の標準化比率 (Standardized Type-Token Ratio) で 40% 程度が標準的な数値とされている（石川 2004）。
- [16] ちなみに、巻末資料にある「テキストプロファイリング・レポート」でとりあげた Sample 4 のテキストの文当たりの平均語数は 23.3 語となっている。これに対して、図 2 で分析した Sample 1 の文当たりの平均語数は 16.4 語である。「標準的」な英文の文当たりの平均語数はおよそ 17 語であり（染谷 1986/1994, Vol. 1, p. 150）、この点でも Sample 4 の難易度は際立っている（平均が 23.3 語ということは、1 文 30 語を超えるような文が相当数含まれているということを示唆するが、実際、Sample 4 の場合は総文数 22 のうち 5 文が 1 文 30 語を超える文になっている）。もちろん、文が長くても達意の文章を書くことは可能であるが、学生にライティングの指導をする場合は、1 文当たりの平均語数が 17 語を大きく超えないように指導すべきであろう。なお、指導に当たっては学生が自ら WLC（またはその他のツール）を使って、自分の書いたものを客観的に自己評価・診断できるように導いていくのが望ましい。
- [17] なお、当然のことながら受験生はすべての教科書を学習しているわけではないことから、平均的な高校生が卒業までに（少なくとも教室内で）学習する語彙の総数はこの基礎 3000 語を相当数下回るものと考えられる（注 6 参照）。
- [18] 問題の総語数も 845 語（全体平均 503 語）で全 40 学部中、上位 2 番目となっている。総語数がかつとも多かったのは都内 M 大学の 937 語であるが、この場合の中高教科書語彙のカバー率は 92.5%、リーダビリティは 5.9 であり、長くても易しい英文が出されていることがわかる。青山学院大学の読解問題は他に比較して長いうえに難しいということになる。

【参考文献】

- Coleman, M. & Liau, T. L. (1975), "A computer readability formula designed for machine scoring," *Journal of Applied Psychology*, Vol. 60, No. 2, pp. 283-284.
- Dale, E., & Chall, J. S. (1948), *A Formula for Predicting Readability*. Columbus, OH: Ohio State University Bureau of Educational Research.
- Flesch, R. (1960), *How to Write, Speak, and Think More Effectively*. New American Library, New York, NY.
- Flesch, R. (1974), *The Art of Readable Writing*. Harper and Row, New York, NY.

- Fry, E.B. (1968), "A Readability Formula That Saves Time," *Journal of Reading*, 11, 7, pp. 265-271.
- Gunning, R. (1952), *The technique of clear writing*. McGraw-Hill International Book Co., New York, NY.
- Gunning, R. (1968), *The Technique of Clear Writing*. McGraw-Hill, New York, NY.
- Impact Information Plain-Language Services (2004), "Robert Gunning's Fog Readability Formula" in *Plain Language At Work Newsletter*, 23 March 2004. [Online] <http://www.impact-information.com/impactinfo/newsletter/plwork08.htm>
- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. Torrance, CA: Laredo Publishing Company, Inc.
- Lauffer, B. (1989), "What Percentage of Text Lexis Is Essential for Comprehension?" In Lauren, C. and Nordman, M. (Eds.) *Special Language: from Humans Thinking to Thinking Machine*, Clevedon: Multilingual Matters, pp. 316-323.
- Lauffer, B. (1992), "How Much Lexis Is Necessary for Reading Comprehension?" In Arnaud and Bejoint (Eds.) *Vocabulary and Applied Linguistics*, London: Macmillan, pp. 126-132.
- Muehleisen, V. (2007), "The SILS Learner Corpus of English." Presentation at the 2007 Summer School of Learner Corpus Research, Centre for English Corpus Linguistics, Université catholique de Louvain. 10-14 September 2007. [Online] <http://www.f.waseda.jp/vicky/learner/index.html>
- Micro Power & Light Co., (N.D.), *Readability Formulas*. [Online] <http://www.micropowerandlight.com/rdformulas.html>
- McLaughlin, G. (1969), "SMOG Grading: A New Readability Formula," *Journal of Reading*, 12, 8, pp. 639-646.
- Nation, I.S.P. (2001), *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Smith, E. A. & Senter, R. J. (1967), "Automated Readability Index," AMRL-TR, 66-22. Wright-Patterson AFB, OH: Aerospace Medical Division, The U.S. Army Defense Technical Information Center (DTIC)
- Taylor, D. (2008), *Readability.Info* [Online] <http://www.readability.info/info.shtml>
- Zakaluk, B.L. & Samuels, S.J. (Eds.) (1988), *Readability: It's Past, Present, & Future*. The International Reading Association, Newark, Delaware.
- 石川慎一郎 (2004) 「日韓の大学入学試験英語問題に見る構成語彙の特徴 英文テキスト・コーパスの解析に基づく考察」 『アジアの英語と英語教育』 7, 1-15. 大学英語教育学会中国四国支部 (2004/9/30)
- 清川英男 (1996) 「リーダビリティ公式とその利用」 『現代英語教育』 9月号, pp. 31-33. 研究社
- 酒井志延 (2006) 「初年次教育・リメディアル教育の現状と課題」 大学基準協会主催 第8回大学評価セミナー講演 (2006年5月17日 於東洋大学白山キャンパス)
- 染谷泰正 (1998) 「AWK による語彙レベル分布計測プログラム Word Level Checker (Ver. 1.5)」 (未刊行オンラインペーパー) [Online] http://www.someya-net.com/kamakuranet/wlc/wlc_manual.html
- 染谷泰正 (1986/1994) 『ライテングマラソン』 (英文ビジネスライターライティング通信講座テキスト: 全4巻) テキスト第3巻, pp. 148-153. アルク (初版 1986, 改訂版 1994)
- 染谷泰正 (2007) 「英語教育における母語の扱いについて メタ言語能力を育てるための "CA + 1" の英語教育の勧め」 *Interactive* 23号, pp. 7-10. 旺文社
- 中條清美・長谷川修治 (2004) 「語彙のカバー率とリーダビリティから見た大学英語入試問題の難易度」 日本大学生産工学部研究報告 B, 2004年6月第37巻, pp. 45-55.

N.D.= Not Dated. オンライン資料のうち、初出または掲載年が明記されていないものを示す。