

# リーダビリティと語彙分析ツールを利用した大学入試長文読解用英文の難易度分析

秦野 進一\*, 倉元 直樹\*\*

\* 東北大学入試センター

\*\* 東北大学大学院教育学研究科 / 東北大学高度教養教育・学生支援機構

**要旨:** 英文の内容に対して感じる質的な難しさの感覚には個人差があり, 尺度化は困難である。一方, 英文の形状的な難しさや使用語彙の難しさは客観的に数値化が可能である。本稿では主として文章の長さや単語の難しさを分析対象として英文の難易度を数値化するリーダビリティ指標と語彙レベルリストに基づいた語彙レベル分析ツールを利用して, 旧七帝大の入試問題で使用された英文を分析する。各大学の長文読解用英文は分量・難易度が様々であり, 英文の内容に応じてどのような試験問題が適しているかも異なってくる。問題作成前の素材文選びの段階でこのような分析を行うことで, 客観的な指標を用いて適切な難易度の英文を選ぶことができると思われる。

**キーワード:** 大学入試, 英語, リーダビリティ, 英文難易度, 語彙分析

## 1. はじめに

受験生の学力に比して難しすぎる試験や逆に易しすぎる試験は識別力が低く, 入学者選抜の資料として十分に機能しない。作題担当者には受験生の学力に合った適切な難易度の試験を安定的に作成することが求められている。常に経験豊富な担当者が作題を担い, 受験生に合った難易度の英文を選ぶことができているなら問題はないが, 必ずしもいつもそのような状況ばかりではないであろう。たとえ経験が浅い担当者が作題を行うことになったとしても, 英文の素材を選択する際に勘や主観に基づくのではなく, 客観的な手法に基づいて行うことができれば適切な難易度の英文を選ぶことができる。今までの研究では入試問題や教科書の英文の分析はリーダビリティや語彙リストに基づいて行われている。例えば Kikuchi (2006) は, 国立・私立の1994年と2004年の個別試験の問題について, 3つのリーダビリティ指標を用いて分析した。その結果, 2つの年の難易度はほぼ同じであり, そのレベルは英語の母語話者にとっても大変難しいレベルであると報告している。また中條・長谷川 (2004) は中学・高校の教科書で習

得する語彙のリストを基に10年分の大学入試センター試験 (以後, 「センター試験」と略記する) の長文読解問題用の英文を分析した。その結果, 年度により差はあるが, 高校卒業時の学力を測定するには適正なレベルであると報告している。これらの研究から英文の難易度分析ツールや語彙リストに基づく語彙分析が英文の難易度を客観的に分析する上で十分機能していることは明らかである。これらを客観的な指標として作題の際に使うことで適切な難易度の素材文を選ぶことができると思われる。

しかし, 入試が多様化している現在, 英語に関わる問題を出題する作題担当者が必ずしも英語教育の知識を豊富に持っているとは限らない。まして, リーダビリティや語彙レベル分析の手法について詳しいとは限らない。そこで本稿では, 英語の試験問題の素材文選びの参考になることを意図し, リーダビリティ指標と語彙レベルリストに基づいた語彙レベル分析ツールの具体的な使用手順について述べる。さらに, それらの手法を用いて旧七帝大の令和4年度入試で使用された英文の難易度分析を行った結果について述べる。

## 2. リーダビリティ指標

英文の難しさを数値化するには、様々なリーダビリティ指標が提案されている。本稿ではワープロソフトの Word の校閲機能を利用することで簡単にリーダビリティが測定できる Flesch Reading Ease と Flesch-Kincaid Grade Level を利用する。

### 2.1 Flesch Reading Ease と Flesch-Kincaid Grade Level

この2種類のリーダビリティ指標は多くの人が利用しているワープロソフトの Word の校閲機能を利用することでインターネットに接続することもなく測定結果が得られる。したがって機密保持のためインターネットを遮断した状態で作題を行わなくてはならない場合でも問題なく使用できる。どちらも英文の平均の長さ（総単語数を英文の数で割った1文あたりの単語数）と1単語あたりの平均音節数を元に公式にあてはめて算出する。1文の長さが長く、単語の平均音節数の多い文章は難易度が高く、1文が短く、単語の平均音節数が短い文章は難易度が低いという考えに基づく指標である。各指標には以下のような特色がある。

#### 2.1.1 Flesch Reading Ease

この指標は以下の計算式によって得ることができる<sup>2)</sup>。なお、日本語訳は第1著者による。

$$206.835 - (1.015 \times 1\text{文あたりの平均単語数}) - (84.6 \times 1\text{語あたりの平均音節数})$$

スコアは0から100までの数値で表され、数が小さいほど英文が難しく、数が大きいほど英文が簡単であることを表している。例えば、英検2級の読解用試験問題の英文では55から65の指標を示している（秦野，2019）。

#### 2.1.2 Flesch-Kincaid Grade Level

上記の指標の公式に改良を加えた以下の計算式で指標が得られる<sup>3)</sup>。

$$(0.39 \times 1\text{文あたりの平均単語数}) + (11.8 \times 1\text{語あたりの平均音節数}) - 15.59$$

この指標は数値がアメリカの学校の児童・生徒の読解レベルを表す学年で表されるので難易度が理解しやすいという利点がある。この指標以外にも The SMOG Index という指標が数値を学年で表している。リーダビリティ指標には様々なものがあるが、それぞれ特有の数値で表されるものが多いため58とか980といった数値がいったいどの程度難しいものなのかは慣れないと理解しにくい。その点この指標は「6」ならば小学校6年生レベル、「9」なら中学3年生レベルの英文であることを表している。日常的にリーダビリティ指標を使い慣れていない人でも難易度をイメージしやすい。

### 2.2 Word を使った測定

Word 2019を利用してリーダビリティの測定を行うための事前設定方法は Microsoft サポート<sup>4)</sup>に図1のように説明されている。

1 [ファイル]、[オプション] の順に移動します。
2 [文章校正] を選択します。
3 [ Word でスペル チェックと文章校正を行う場合] で、[スペル チェックで文章校正を行う] チェックボックスを オンにします。
4 [文書の読みやすさを評価する] を選択する

図1. リーダビリティ指標の事前設定方法

この設定を行った後、測定対象のファイルを開き、F7キーを押してスペルチェックを実行するか、「校閲」→「スペルチェックと文章校正」をクリックすると表記上の修正候補などが提示され、文章校正が終了すると読みやすさの評価数値が出てくる。一番下の Readability の欄に Flesch Reading Ease と Flesch-Kincaid Grade Level の数値が表示される。通常は自動的に日本語と英語を判断して分析が行われるが、英文の分析を行ったのにリーダビリティ指標が表示されない場合には、上記事前設定方法のファイルオプションで「言語」を選択して校正機能で優先する言語を英語に設定するとよい。

### 2.3 旧七帝大の長文読解用英文の難易度分析

旧七帝大の令和4年度前期入試で使用された英文を以下の手順にしたがって準備し、上記2つの

リーダビリティ指標を用いて分析を行った結果を表1にまとめた。なお本稿では各大学の長文読解問題のうち、最も語数の多い大問の本文についてのみ分析を行った。ただし大阪大学については外国語学部（以下、「外国」と略記）とそれ以外の学部（以下、「外国以外」と略記）で一部使用する長文読解問題が異なるので、それぞれ最も語数の多い問題について分析を行った。

- (1) 大学入試問題のデータベースである Xam (イグザム) 英語2022<sup>5)</sup> から旧七帝大の長文読解問題を取り出す。問題は Word か一太郎のいずれでも出せるので Word を使用する。その中から今回の分析対象となる最も語数の多い問題を選ぶ。電子データの無い英文を分析する場合には、OCR ソフト（スキャナーで読み取った画像データをテキストデータに変換する文字認識ソフトウェア）を利用して新聞や雑誌の記事などをテキストに変換して Word に張り付ける。OCR ソフトは多くの種類が市販されているが無料で使えるフリーソフトもある。
- (2) 本文として使用されている英文の下線、記号、番号等、問題作成者によって付加されたものを取り除き、また空欄に単語、語句、文章などを補充する問題がある場合には、空欄に正解を入れて加工前の元の英文に復元する。
- (3) 2.2で述べた手順にしたがってリーダビリティの測定を行う。

表1を見ると各大学の読解問題で使用された英文の分量・難易度にはかなりの違いがあることが

わかる。あくまで今回分析の対象とした問題に限定した議論ではあるが、8題の平均語数は837語で、大阪大（外国）（1071語）が一番多く、東北大（985語）、東京大（977語）が次いで多い。一番少ない京都大（595語）と一番多い大阪大（外国）では476語の差がある。

分析対象とした8題の中で最も難易度が高かった（Flesch Reading Ease の数値が小さく、Flesch-Kincaid Grade Level の数値が大きいものを指す）のは京都大の問題で、2つの指標の双方で8題中最も難しかったことが示された。Flesch-Kincaid Grade Level の数値は、アメリカの大学2年生が読むレベルの英文であることを示している。次いで東北大（高校3年生レベル）、九州大（高校2年生レベル）が難易度の高い英文となっている。最も難易度が低かったのが東京大の問題で、小学校6年生レベルである。ただし、東京大は長文読解問題を5題出題しているため、今回対象となった「最も語数の多い英文」に関しては平易なものが使用されているということを示しているに過ぎない。1つの英文に関して、何を問い、どのように答えさせるかによって問題は易しくもなり、難しくもなるので、ここにあげた数値だけをもって問題が平易であるということとはできない。リーダビリティの数値はあくまで使用されている英文の難易度を示すものである。

#### 2.4 その他のリーダビリティ

本文として利用された素材文のおおよその難易度を知るには上記のいずれかの指標を用いれば十分だと思われるが、もっと多くの指標を参考にしたいという場合には、Readability Formulas<sup>6)</sup> というウェブサイトにある Automatic Readability

表1 旧七帝大の長文読解用英文分析結果

大学名	北海道大	東北大	東京大	名古屋大	京都大	大阪大		九州大	平均	
						外国語学部	外国語以外			
長文読解用英文の数	4	2	5	2	2	2	2	3	2.8	
分析対象大問番号	1	1	5	1	2	2	2	2		
語数	862	985	977	786	595	1071	719	701	837.0	
リーダビリティ	Flesch Reading Ease	57.8	44.2	79.3	54.9	34.4	59.9	66.2	47.9	55.6
	Flesch-Kincaid Grade Level	9.1	12.6	6.1	11.8	14.1	9.5	8.2	11.8	10.4

Checker<sup>7)</sup>を利用して分析したい英文を送れば、同時に7種類のリーダビリティ指標を返してくれる。研究目的でなければここまで多くの指標は不要なので、作題担当者に合った使いやすいのものを選べばよい。

また、米国 MetaMetrics 社が開発した Lexile measures という指標は、英文の難しさを表すと同時に読み手の読解力も表す指標となっている。Lexile reading measure と Lexile text measure<sup>8)</sup>の2種類あり、前者はテストの結果などによって計測された受験者の読解力を数値で示し、後者は「1文あたりの長さ」と「単語の出現頻度」から分析した単語の難易度によって文章の難易度を数値で示す。そのため教育現場では生徒の読解力に合わせたテキストを選択する際などに使われている。これを応用すれば例えば授業を担当するクラスの学生の Lexile reading measure を計測して学生に適した難易度の英文を選ぶといった利用法も可能である。TOEFL のウェブサイトでは子どもたちがこの指数を利用して自分の英語力に合った洋書を選ぶようになっている<sup>9)</sup>。

リーダビリティは主として英文の形状的な難易度を利用して測定しているのだから、実際に英文を読む対象者にとって本当にその数値が示す難易度通りに難しく感じるかどうかは、人間の目で英文を読んで確かめる必要がある。あくまで参考資料として利用することをお勧めする。

### 3. 単語の難易度分析

我々が英文を読むときに、その英文を「難しい」と思う大きな要因の1つは単語の難しさである。知らない単語が多く出てくる英文を読むことはつらく、辞書に頼ることなく内容を理解することが難しいことも多いであろう。作題者は受験生にある程度の語彙力を持っていることを期待して素材文を選ぶであろうが、受験生の英語力に比して難しすぎる英文では英語力を適切に測ることはできない。受験生の英語力を測るのに適した難易度の英文を選ぶためには使用されている単語の難易度を理解した上で素材文を選ぶことが大切である。単語の難易度は、しかるべき語彙レベルリストに基づいて使用されている単語について分析をし、難易度の高い語がどの程度含まれているかをみる

ことで判断できる。本稿では日本人英語学習者を対象とした単語の分析に適していると思われる語彙レベルリストである新 JACET 8000 と CEFR-J Wordlist、それに単語レベルチェッカー（語彙レベル分析ソフト）<sup>10)</sup>を利用して語彙レベル分析を行う。

#### 3.1 新 JACET 8000

新 JACET 8000（大学英語教育学会基本語改定特別委員会、2016）は2005年に大学英語教育学会により日本人英語学習者のための教育語彙集として出版された8000語の英単語リスト「JACET 8000」（相澤ほか、2005）の2016年改訂版である。旧版は BNC（The British National Corpus）という1990年前後に作られた大規模なイギリス英語コーパスを母体に編集されたが、新版では BNC に加えて、アメリカ英語のコーパスである COCA（The Corpus of Contemporary American English）を母体としてベースとなるリストを作成し、さらに日本人英語学習者の英語学習の実態および英語学習の目標に合わせたものに補正するための資料を基に補正された。資料として使われたのは、中学・高校の検定教科書、大学入試センター試験問題、47都道府県の公立高校入試問題、英検・TOEFL・TOEIC の問題、日本の英字新聞、英語による学術入門書などである。

JACET 8000は1000語単位でレベル分けがされており、2005年版の目次では表2のようにおおそのレベルについての説明が掲載されている。また、新 JACET 8000には本体としてのリスト以外に「中学・高校コミュニケーション支援語彙リスト」、「共通学術語彙リスト」、「発信語彙リスト」の3つの付加リストも加えられている。

令和2年度から小学校で全面実施されている新しい学習指導要領（文部科学省、2017）では小学校でも英語の教科化が始まり、600語から700語の単語学習が行われている。さらに中学校、高校でも学習する単語数は増えているので、小学校、中学校、高校を通した総学習単語数は旧学習指導要領では3000語だったのが新しい学習指導要領では4,000語から5,000語へと大幅に増えることになる。現在では上記リストの Level 3が「高校英語教科書レベルの単語」であるが、新学習指導要領が完全

実施された際には Level 4 もしくは Level 5 までの単語が「高校英語教科書レベルの単語」となる。

表2 JACET 8000(2005)のレベル分類

レベル	順位	説明
Level 1	1-1000	中学校の教科書に頻出する基本単語
Level 2	1001-2000	高校初級レベルの単語英検準2級レベル。
Level 3	2001-3000	高校英語教科書レベルの単語、センター試験。
Level 4	3001-4000	大学受験、及び大学一般教養の初級レベルに相当。
Level 5	4001-5000	難関大学受験、及び大学一般教養に相当。
Level 6	5001-6000	英語を専門としない大学生やビジネスマンが目指すレベル。英検準1級レベル。
Level 7	6001-7000	英語専攻の大学生や仕事で英語を使うビジネスマンの到達目標とするレベル。英検1級, TOEICの95%以上の単語をカバーしている。
Level 8	7001-8000	日本人の英語学習者一般的な単語学習の最終到達目標。

仮に、この分類で Level 4 レベルまでの単語力(項目数累計4,000語)を受験生に求めるならば、Level 5 レベル以上の語について注を付けるかどうかを検討すればよい。また選抜性の高い大学であれば Level 5 までの5,000語を受験生に求めることとして Level 6 以上の語について注を付けるかどうか、検討することになる。前後関係から意味を推測できる語や、すでに日本語として使われている語などに対する注は不要であるので、設定したレベル以上の語だからといって機械的に注を付ける必要はない。

### 3.2 CEFR-J Wordlist

CEFR-J のウェブサイト<sup>11)</sup>によれば、「CEFR-J は欧州共通言語参照枠(CEFR)をベースに、日本の英語教育での利用を目的に構築された、新しい英語能力の到達度指標」である。指標自体は「聞くこと」、「読むこと」、「書くこと」、「話すこと(やりとり)」、「話すこと(発表)」の5技能それぞれについて12レベルの「言葉を使って何ができる

か」を示す CAN-DO リストの形式を取っている。例えば B1.1 レベルの「話すこと(やりとり)」では、「個人的に関心のある具体的なトピックについて、簡単な英語を多様に用いて、社交的な会話を続けることができる」、B1.2 レベルの「聞くこと」では、「はっきりとなじみのある発音で話されれば、身近なトピックの短いラジオニュースなどを聞いて、要点を理解することができる」などと表記されている。この CEFR-J に付属する資料として CEFR-J Wordlist がある。約7,000語が A1 から B2 までの4レベルに分類されている。おおよその相当レベルを表3に示す。

表3 CEFR-J Wordlistのレベル分類

レベル	単語数	日本の学校教育での相当レベル
A1	1068	小学校～中学1年程度
A2	1352	中学2年～高校1年程度
B1	2353	高校2年～大学受験レベル
B2	2692	大学受験～大学教養レベル
合計	6868	

仮にこの分類で B2 レベルまでの単語力(項目数累計6,868語)を受験生に求めるならば、(B2 までのリストに)該当なし(NA: Not Applicable)と表示された単語について注を付けるかどうかを検討すればよい。

### 3.3 単語レベルチェッカー

単語レベルチェッカーはイー・キャスト社製の市販ソフトで、中学校・高校の各教科書の持つシェアのパーセンテージに基づいて生徒の何%がその単語を知っているかという「認知率」を学年別に数値化し、さらに過去10年分のセンター試験の本試験の全単語を分析して、6段階のレベルに分けた8,517語のリストに基づいて語彙レベルのチェックを行うものである。自分のパソコンにインストールして使用するので、インターネットにつながることなく分析が可能である。Wordでのリーダビリティ分析同様、機密保持の観点からインターネットを遮断して作題しなくてはならない場合にも使用できる。毎年新しい試験のデータを元に改訂版を作成しているので、最新の単語などもリストに反映されている。分析後にはレベル外単語数、レベル外単語率などが表示され、難語を平

易語への置き換えをサポートするシソーラス（類義語辞典）機能も搭載されている。レベルのおおよその目安を表4に示す。なお現在、大学入学共通テストの開始に伴い、新バージョンへのプログラムの改修中であり、バージョン2020が最新版となっている。

仮にこの分類で「高3 / センター試験」レベルまでの単語力（項目数累計8,517語）を受験生に求めるならば、レベル外と示された単語について注を付けるかどうかを検討すればよい。

表4 単語レベルチェッカーのレベル分類

レベル	単語数	基準
中1	482	中1の認知率50%以上
中2	419	中2の認知率45%以上
中3	445	中3の認知率35%以上
高1	965	高1の認知率25%以上
高2	1588	高2の認知率15%以上
高3 / センター試験	4618	高2の認知率14.9%以下、または高3新出語、センター試験
合計	8517	

### 3.4 語彙分析の具体的方法

以下に、それぞれの語彙リストに基づいた分析を行う具体的な方法を示す。

#### 3.4.1 New Word Level Checker<sup>12)</sup> 利用の語彙分析

オンラインで利用できる英文語彙難易度解析プログラムである New Word Level Checker を利用すれば上記の新 JACET 8000 と CEFR-J の語彙リストに基づいた難易度分析ができる。このサイトでは他にも SVL 12000<sup>13)</sup> などの5種類の語彙リストに基づく語彙分析が可能である。利用方法はいたって簡単で、ウェブサイトに表示されているテキストボックスに分析したい英文を入れ、使用したい語彙リストを選択して「Check」ボタンをクリックするだけで語彙レベルごとにテキストカバー率とそのグラフ、難易度ごとに色分けしたテキストを返してくれる。その後「Word List」をクリックすればすべての語のレベルの一覧表が返ってくる。この Word List はボタン1つで Excel ファイルとしてダウンロードできるので、レベル順に並び替えて難しい語の一覧を表示させるなどの使

い方ができる。4色～5色で色分けされたテキストの表示は色によってはやや見分けがつきにくいという難点がある。その場合にはダウンロードした Word List の一覧表を併用して注を付けるかどうか判断するとよい。

New Word Level Checker では固有名詞や数字は既知語として分類される (PropNoun\_Num と表示) ので、それ以外の語についてそれぞれの語彙レベルが表示される。

また、reason という語は「理由」という意味の名詞で使用されれば平易な語だが、「推論する、説得する」という意味の動詞になると難しくなる。CEFR-J ではこのような品詞まで含めて分析しているので、black (黒くなる、黒くする：動詞) なども難語に分類されている。

ほかにも、原形と過去形をそれぞれリストアップするかどうか、短縮形をどう扱うかなどは使用する語彙リストによって異なる。

#### 3.4.2 単語レベルチェッカー利用の語彙分析

インストールした単語レベルチェッカーを開き、テキストボックスに分析対象の英文をコピー＆ペーストする。「高2」「高3 / センター試験」などのレベルを設定した上でレベル判定ボタンを押せば判定結果が表示される。

設定したレベルを超える難易度の語（難語）が赤色、それ以外の語は黒字で表示されるので難語が一目でわかり見やすい。前述の2つのリストでは難語に分類されてしまうアルファベット単体や記号は単語レベルチェッカーでは難語には分類されていない。

いずれの語彙レベルリストを用いても、ときどきどうしてこの語が難語と分類されているのだろうか（あるいは、その逆もある）ということがあるので、分析結果をそのまま鵜呑みにするのではなく、どのような語が難語と分析されたのかをリストで確認することを勧める。

#### 3.4.3 旧七帝大の長文読解用英文の語彙分析

2.3で用意した旧七帝大の令和4年度前期試験で使用された長文読解用英文をコピーして、3.4.1と3.4.2で示したテキストボックスに入れて語彙分析を行った結果をまとめたものが表5である。設

定した語彙レベルはそれぞれ Level 6 (新 JACET 8000), B2 (CEFR-J), 高3/ センター試験 (単語レベルチェッカー) である。CEFR-J ではリスト外 (NA) に分類された語からアルファベット単体 (A や X など) や記号 (% など) を除いた語を難語と定義し, 新 Jacet8000 ではそれに Level 7 以上の語を加えたものを難語として扱った。単語レベルチェッカーではレベル外単語と表示された語を難語として扱った。

平均難語数が一番多かったのは大阪大 (外国: 41.7語) で, 最も少ない京都大 (18.7語)・九州大 (18.7語) とは23語の差がある。次に多かったのは東北大 (34.3語) と大阪大 (外国以外: 32.3語) であるが, それぞれ注語として東北大は8語, 大阪大 (外国以外) は3語に説明をつけている。東京大は平均難語数が31.3語であるが, 注語として10語に説明をつけているので実際は難語数としては20語程度である。難語率で比較すると3%台前半の大学が多い。大阪大 (外国以外) が4.5%とやや高く, 九州大が2.7%とやや低い。

#### 4. 各大学の問題文の特長

表1のリーダビリティと表5の語彙レベルの2つの分析結果から, 各大学の問題文には以下のような特徴があることがわかる。以下, 構文難易度の高い順に記載する。

京都大: 英文量は多くないが, 構文の難しい (1文の長さが長い) 英文である。難語数は多くなく, 難語率はそれほど高くないので, 語彙力よりは構文理解力を問うのに適した英文で

ある。

東北大: 英文量が多い。構文難易度も高く, 難語も多く, 難語率も比較的高い。語彙力, 構文理解力ともに要求される英文である。

名古屋大: 8題の中では英文量はさほど多くはないが, 構文の難易度は高めである。難語率は高くない。

九州大: 英文量はあまり多くない。構文の難易度は高めである。難語は少なく, 難語率も低い。

大阪大 (外国): 英文量が多く, 構文難易度は8題の中では平均的な難しさである。難語数も多く, 難語率も高い一方で注語が1つもない。かなりの語彙力が必要とされる英文である。

北海道大: 英文量は8題のほぼ平均であり, 構文の難易度はやや平易, 難語数, 難語率も8大学のほぼ平均である。

大阪大 (外国以外): 英文量はさほど多くなく, 構文もそれほど難しくはないが難語数が多く, 難語率が高い。語彙力が必要とされる英文である。

東京大: 英文量は多いが, 構文の難易度が低いので英文の意味を理解することは難しくない。英文を訳す, あるいは構文の理解度を問うような問題では平易になりすぎてしまうので適さない。

#### 5. 素材文選択時の活用例

本稿で分析したのはすでに実施された試験問題の英文難易度だが, 同様の分析を作題前の素材候

表5 旧七帝大の長文読解用英文語彙分析結果

	北海道大	東北大	東京大	名古屋大	京都大	大阪大		九州大	平均
						外国語学部	外国語以外		
語数	862	985	977	786	595	1071	719	701	837.0
JACET	28	15	27	18	17	43	29	14	23.9
CEFR-J	29	51	37	32	20	51	44	25	36.1
単語レベルチェッカー	26	37	30	23	19	31	24	17	25.9
平均難語数	27.7	34.3	31.3	24.3	18.7	41.7	32.3	18.7	28.6
難語率	3.2%	3.5%	3.2%	3.1%	3.1%	3.9%	4.5%	2.7%	3.7%
注語数	1	8	10	0	0	0	3	2	3.0

補の英文で行うことで客観的な指標を用いての素材選択が可能になる。例えば、難しい構文の英文の意味を正しく理解できるかどうかを見たい場合、京都大、東北大、九州大、名古屋大などの英文のようにリーダビリティが難解であることを示している英文が適している。それに加えて高度な語彙力も身につけているかどうか見たい場合には東北大のように難語率の高い英文が適している。

分量の多い英文を与えて必要な情報を読み取らせるような問題は、東大、大阪大(外国)のように語数が多く、リーダビリティもそれほど高くない英文が適している。さらに、ある程度の語彙力があることを受験生に求める場合には大阪大(外国)のように語彙レベルが難解であることを示し、平均難語数が多い英文が適している。

しっかりとした語彙力を身につけた学生を選抜したい場合、大阪大(外国、外国以外)のような難語の多い英文を使用し、注語も必要最小限にするという出題の仕方が考えられる。

## 6. おわりに

本稿で紹介したリーダビリティや語彙レベル分析ツールは、慣れないと作業が煩雑に思えるかもしれない。また、得られた数値もどう解釈しているのかよく分からないこともあるであろう。しかし、どれも慣れれば5分程度で終わるものばかりである。作題担当者にとって理解しやすいと思われる指標を継続的に使っていれば、慣れるのにそれほど長い時間はかからない。そして、経年で同じ指標を用いて分析していくとともに、試験結果も合わせて分析すれば受験生の英語力を測定するのに適切と思える難易度の英文がどの程度のもか見えてくる。初めは、過去に出題した問題の中でうまくできていたと思われる英文を分析して、その難易度に近いものを選ぶ、という方法も可能である。今回紹介した分析ツールが客観的な指標に基づいた素材文選択に活用されることを期待している。

## 注

- 1) 北海道大学、東北大学、東京大学、名古屋大学、京都大学、大阪大学、九州大学。
- 2) <https://www.readabilityformulas.com/>

[flesch-reading-ease-readability-formula.php](https://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php) (2022.11.21)

- 3) <https://readabilityformulas.com/flesch-grade-level-readability-formula.php> (2022.11.21)
- 4) [https://support.microsoft.com/ja-jp/office/%E3%83%89%E3%82%AD%E3%83%A5%E3%83%A1%E3%83%B3%E3%83%88%E3%81%AE%E8%AA%AD%E3%81%BF%E3%82%84%E3%81%99%E3%81%95%E3%81%A8%E3%83%AC%E3%83%99%E3%83%AB%E3%81%AE%E7%B5%B1%E8%A8%88%E6%83%85%E5%A0%B1%E3%82%92%E5%8F%96%E5%BE%97%E3%81%99%E3%82%8B-85b4969e-e80a-4777-8dd3-f7fc3c8b3fd2#ID0EBDD=Older\\_versions\\_of\\_Office](https://support.microsoft.com/ja-jp/office/%E3%83%89%E3%82%AD%E3%83%A5%E3%83%A1%E3%83%B3%E3%83%88%E3%81%AE%E8%AA%AD%E3%81%BF%E3%82%84%E3%81%99%E3%81%95%E3%81%A8%E3%83%AC%E3%83%99%E3%83%AB%E3%81%AE%E7%B5%B1%E8%A8%88%E6%83%85%E5%A0%B1%E3%82%92%E5%8F%96%E5%BE%97%E3%81%99%E3%82%8B-85b4969e-e80a-4777-8dd3-f7fc3c8b3fd2#ID0EBDD=Older_versions_of_Office) (2022.11.21)
- 5) イグザム2022英語 株式会社ジェイシー教育研究所
- 6) <https://readabilityformulas.com/> (2022.9.8)
- 7) <https://readabilityformulas.com/free-readability-formula-tests.php> (2022.9.8)
- 8) <https://hub.lexile.com/analyzer> (2022.11.21)
- 9) <https://toefljunior.lexile.com/ja/> (2022.11.21)
- 10) 単語レベルチェッカー 2020 イー・キャスト
- 11) <http://cefr-j.org/cefrj.html> (2022.9.8)
- 12) 関西大学大学院外国語教育研究大学院応用言語学部の水本敦教授によって開発されたプログラムで、前身のプログラムは青山学院大学文学部英米文学科の染谷泰正教授が開発した Word Level Checker である。 <https://nwlc.pythonanywhere.com/> (2022.11.21)
- 13) 株式会社アルクが独自に開発した語彙リストで、最上級のレベルの難語を含む12,000語を1000語ずつの12レベルに分類している。

## 謝辞

本研究は JSPS 科研費 JP20K20421 の助成を受けたものである。

## 文献

相澤一美・石川慎一郎・村田年編集代表 デビッド・クルソン英文校閲(2005).『「大学英語教育学会基本語リスト」に基づく JACET8000英単語』桐原書店。



- 大学英語教育学会基本語改定特別委員会 (2016). 『大学英語教育学会基本語リスト 新 JACET8000』 桐原書店.
- 秦野進一 (2019). 「資格・検定試験における長文読解用英文の難易度比較」『大学入試研究ジャーナル』, 29, 176-182.
- Keita Kikuchi (2006). “Revisiting English Entrance Examinations at Japanese Universities after a Decade” JALT Journal, Vol.28, No.1, 77-96.
- 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』, 37, 45-55.
- 染谷泰正 (2009). 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, 『青山学院大学文学部紀要』 51, 99-122.

## **Analysis of the Difficulty Level of English Sentences for Entrance Examinations Using Readability and Lexical Analysis**

**Shinichi HATANO\***, **Naoki T. KURAMOTO\*\***

\* Admission Center, Tohoku University

\*\* Graduate School of Education / Institute for Excellence in Higher Education, Tohoku University

### **ABSTRACT**

The qualitative difficulty felt with the content of English texts varies from person to person and is difficult to quantify. On the other hand, the difficulty of a particular English sentence structure and the difficulty of the vocabulary used can be objectively quantified. In this paper, we analyze the English sentences used in the entrance examination of the former Seven Imperial Universities using a vocabulary-level analysis tool based on a readability index and a vocabulary level list that quantifies the difficulty of English sentences mainly by analyzing sentence length and word difficulty. The amount and difficulty of English texts used for reading comprehension tests vary by university, and the type of test questions that are appropriate for each type of English text will differ depending on the content of the English text. By performing such analysis at the stage of selecting the texts, before preparing the questions, it will be possible to select English texts that are suitable for measuring the English ability of examinees.

**Key words:** university entrance examinations, English, readability, difficulty level of English sentences, vocabulary-level analysis