

【原著】

学習者向け英字新聞 2 紙における CEFR-J の有効性の比較

八 島 等

A Comparison between Two English Newspapers for Learners
with Regards to the Effectiveness of the CEFR-J Wordlist

Hitoshi Yashima

4 年前から、CEFR-J で示されている A1 から B2 までの語彙が一般紙の英字新聞でどの程度用いられているのかを 3 年間調査した。その結果、CEFR-J の収録語以外の語が約 14-18% もあり、CEFR-J の有効性に疑問が残った。そこで、昨年、研究対象を学習者向け英字新聞に変更して同様の調査をした。すると、A1 が約 29-32%、A2 が約 22-23%、B1 が約 20-24%、B2 が約 12-13%、それ以外の語が約 10-12% であり、やはり CEFR-J の有効性に疑問が残ることがわかった。しかし、CEFR-J の語の出現頻度を調査すると、約 92-94% であり、「最低限の」語彙カバー率に近いものであった。今回は別の学習者向け英字新聞を調査したところ、A1 が約 28-29%、A2 が約 21-23%、B1 が約 22-23%、B2 が約 13-14%、それ以外の語が約 13-14% となった。また、出現頻度は約 93-94% で、ほぼ昨年同様の結果となった。出現頻度の場合、CEFR-J 収録語以外の語の中に品詞違い語や派生語が含まれている。昨年の調査では約 32-37%、今年の調査では約 33-35% それぞれ含まれているので、実際の CEFR-J の語の出現頻度は、いずれも約 95-96% 以上となり、「最低限の」語彙カバー率であると言える。

キーワード：CEFR-J, 学習者向け英字新聞, 有効性, 比較

研究動機

10 年前に、日本人英語学習者（高校生）にとって文脈からの意味推測のために「最適な」未知語の割合を求めて、50語、60語、80語、100語に 1 語という 4 つの割合で調査した。その結果、100語に 1 語（99% の語彙カバー率）の場合、語彙サイズが 3,000 語（word family）程度あれば、平均で 60% 以上の、4,000 語程度あれば、平均で 80% 以上の、それぞれ推測成功率が得られることがわかった。また、80語に 1 語（98.75% の語彙カバー率）の場合でも、語彙サイズが 4,000 語程度あれば、推測成功率は平均で 80% 以上であることがわかった（八島, 2013）。3,000-5,000 語は「一生懸命努力すれば、学習者にとって達成可能であるように思われる」（Schmitt, Jiang, & Grabe, 2011）語数である。以上のことから、日本人英語学習者にとって「最適な」未知語の割合は、100語に 1 語（語彙カバー率 99%）であることがわかった。

また、8 年前、読解に関して、98% 以上の語彙カバー率の題材の内容を、どの程度理解することができるのかを調査した。その結果、学習者には、98% の語彙カバー率は「最適な」ものとは言えないことが判明した。語彙カバー率が 99% の場合のみ、平均で 60% 以上の読解問題の正答率を得ることができた（八島, 2015）。

以上の2つの研究から、日本人英語学習者にとっては、99%の語彙カバー率が必要であると考えられる。この99%という値は、Carver (1994) を支持するものと言える。Carver (1994) は、小学生及び大学院生の英語母語話者を対象とした研究により、題材が比較的易しい場合、未知語はほぼ0%であり、題材が比較的難しい場合、2%以上が未知語であり、題材の難易度が読み手の能力にはほぼ一致している場合には、約1%が未知語であると指摘している。このことを語彙カバー率に当てはめてみると、母語の場合でも、99%が適切な語彙カバー率であると考えられる。従って、日本人英語学習者にとっても当てはまるということは当然であろう。

では実際に、日本人英語学習者はその99%の語彙カバー率に達するために、どのような語を何語程度習得すればよいのであろうか。そこで参考となるのが、最近注目を集めている欧州評議会 (Council of Europe) が2001年に公表した言語能力の到達度指標である Common European Framework of Reference for Languages: Learning, teaching, assessment (学習, 教授, 評価のためのヨーロッパ言語共通参照枠), 略称, CEFR である。英語に限らず、外国語に関する言語能力を A (基礎段階の言語使用者), B (自立した言語使用者), C (熟達した言語使用者) の3つのレベルに大きく分け、それをさらに、A1・A2, B1・B2, C1・C2の合計6段階のレベルに分けて表している。

そのCEFRを日本に導入するために、構築されたのがCEFR-Jである。そのプロジェクトの中で、CEFR-J Wordlist が作成され、現在は Version 1.6 で、A1は1166語、A2は1411語、B1は2445語、B2は2779語の合計7801語が収められている (投野 (2020) では「語」ではなく、「項目」となっている)。

それでは、日本人英語学習者 (高校生) の英語学習の基盤となる検定教科書の中で、CEFR-J に収められている語がどの程度用いられているのかということが疑問となり、3年かけて、改訂版の1～3年生用の上級の検定教科書から採択部数の多い3冊を選び調査した。その結果、レベル別分布状況は、1冊ずつでは、A1が約17-24%, A2が約16-20%, B1が約19-22%, B2が約11-13%, それ以外の語が約23-32%であり、3冊の合計では、A1が13.39%, A2が13.69%, B1が20.01%, B2が14.03%, それ以外の語が38.88%であった。3冊の合計となると、A1の収録語が少ないので、全体における割合は少なくなってしまうが、1冊ずつで見ると、多い順に、B1, A1, A2, B2という分布状況であった (1冊のみ A1, B1, A2, B2の順)。また、レベル別カバー率は、1冊ずつでは、A1が約79-86%, A2が約57-64%, B1が約34-49%, B2が約16-25%であり、3冊の合計では、A1が93.56%, A2が78.74%, B1が66.41%, B2が40.96%であった。検定教科書には、A2が予想外に出現しないことがわかった (八島, 2019)。

検定教科書での研究の延長線上に位置する研究として、英字新聞 *The Japan Times* の一週間分 (日曜版は形式が違うので研究対象から除外) の中で、CEFR-J に収められている語がどの程度用いられているのかという調査を総異なり語のレベル別分布状況とレベル別出現頻度の観点から行った。その結果、レベル別分布状況は、一日分ごとに、A1が約20-22%, A2が約19-21%, B1が約25%, B2が約17-18%, それ以外の語が約14-18%であった。一日分ごとに見ると、多い順に、B1, A1, A2, B2という分布状況であった (ある一日のみ、B1, A2, A1, B2の順)。この結果は、上記の検定教科書の分布状況と同じであることがわかった。また、レベル別出現頻度は、一日分ごとに、A1が約60-61%, A2が約13-14%, B1が約11-12%, B2が約6-7%, それ以外の語が約8%であった。レベル別分布状況の観点からみると、CEFR-J の英字新聞に対する有効性に疑問が残った。しかし、レベル別出現頻度の観点からみると、92%のカバー率となり、さらに、すべてのそれ以外の語のうちの35.74%がCEFR-J の収録語の品詞違い語や派生語なので、実際のCEFR-J の収録語以外の語の出現率は5.14%以下となり、Laufer & Ravenhorst-Kalovski (2010) が提唱する

「最低限の」(minimal) 閾値である95%とほぼ一致することがわかった(八島, 2021a, b)。

そこで、昨年、学習者向け英字新聞 *The Japan Times Alpha* の一か月分を研究対象として上記の一般紙と同様の調査をした。すると、レベル別分布状況は、一日分平均で、A1が約29-32%、A2が約22-23%、B1が約20-24%、B2が約12-13%、それ以外の語が約10-12%であった。学習者にとって、学習者向けは一般紙よりは読みやすいが、やはり CEFR-J の有効性に疑問が残ることがわかった。しかし、CEFR-J の語の出現頻度は、一日分平均で、約92-94%であり、「最低限の」語彙カバー率に近いものであった。さらに、すべてのそれ以外の語のうちの32-37%が CEFR-J の収録語の品詞違い語や派生語であったので、実際の CEFR-J の収録語以外の語の出現率は約4.0-5.3%以下となり、「最低限の」閾値である95%とほぼ一致することがわかった(八島, 2022a, b, c)。

研究目的

本研究の主な目的は、CEFR-J に収められている語がどの程度、昨年とは別の学習者向け英字新聞で用いられているのかを調査することである。上記の通り、学習者向け英字新聞に対する CEFR-J の有効性に疑問を感じたために、別の学習者向け英字新聞における CEFR-J の収録語のレベル別分布状況と CEFR-J の収録語以外の語の出現率を調査して、CEFR-J の有効性と CEFR-J の収録語以外の語の出現率を調査して、昨年の調査結果との比較をする。

研究方法

使用テキスト

使用したのは許可を得てダウンロードをした東京外国語大学投野由紀夫研究室の『CEFR-J Wordlist Version 1.6』と購入した *Asahi Weekly* の2022年11月6, 13, 20, 27日版である。

CEFR-J Version 1.6 には上記の通り、A1~B2までの7801語が収められている。その7801語と学習者向け英字新聞で用いられている語との対応を調査した。本研究では、CEFR-J の特徴に従って、固有名詞は月や曜日など一部を除いて研究対象から除外した。

研究手順及びデータ分析の方法

学習者向け英字新聞をスキャナーで取り込み、Web で公開されている、関西大学の水本篤先生が作成された New Word Level Checker にかけた。その上で、再度、CEFR-J Wordlist Version 1.6 と対照し、1語1語レベルの確認をして、品詞違い語、派生語の選定を行った。但し、固有名詞は月や曜日など一部を除いて入力対象から除外した。CEFR-J の A1, A2, B1, B2 の各レベルにおいて、学習者向け英字新聞では何語用いられているのか、また、CEFR-J に載っていない語が何語あるのかを調べ、総異なり語のレベル別分布状況を調査した。さらに、総異なり語のレベル別出現頻度を調査した。出現頻度の調査においては、CEFR-J の収録語以外の語の中に、CEFR-J の収録語の品詞違い語や派生語が含まれる。この分布状況と出現頻度という2つの観点から、CEFR-J の有効性を検証して、昨年の調査結果と比較した。

結果

学習者向け英字新聞の使用語彙の CEFR-J のレベル別分布状況

A1からB2及びそれ以外の語の分布状況を見ると、4日分の学習者向け英字新聞のいずれも、

似たような状況になっている。すなわち、A1が約28-29%、A2が約21-23%、B1が約22-23%、B2が約13-14%、それ以外の語が約13-14%をそれぞれ占めている。このことから、4日分の各レベルが占める割合の違いは最大でも約2%にとどまっていることがわかり、記事内容の違いにかかわらず、レベル別分布状況の違いはほとんど生じないことがわかった。また、A1からB1までで約72-74%と全体の4分の3近くを占めていることもわかった。さらに、割合の多い順を見ると、傾向が前半2日分と後半2日分に分かれ、11月6日分はA1、A2、B1、それ以外の語、B2、11月13日分はA1、A2、B1、B2、それ以外の語という順序にそれぞれなっており、11月20日、27日分は共に、A1、B1、A2、それ以外の語、B2という順序になっていることもわかった。A2とB1、B2とそれ以外の語の順序が日によって違っているが、その差は多くても2%程度であることから、各レベルが占める割合の順序の違いもそれほどないことがわかった（表1～4）。

表 1
Asahi Weekly 11月6日分の CEFR-J のレベル別分布状況

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	687語	20語	15語	722語	27.62
A2	547語	25語	18語	590語	22.57
B1	516語	16語	49語	581語	22.23
B2	312語	19語	29語	360語	13.77
それ以外の語				361語	13.81
合計	2,062語	80語	111語	2,614語	

表 2
Asahi Weekly 11月13日分の CEFR-J のレベル別分布状況

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	662語	13語	14語	689語	28.55
A2	518語	19語	17語	554語	22.96
B1	493語	10語	28語	531語	22.01
B2	287語	18語	30語	335語	13.88
それ以外の語				304語	12.60
合計	1,960語	60語	89語	2,413語	

表 3
Asahi Weekly 11月20日分の CEFR-J のレベル別分布状況

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	658語	14語	18語	690語	27.92
A2	480語	15語	20語	515語	20.88
B1	507語	19語	43語	569語	23.03
B2	284語	25語	34語	343語	13.88
それ以外の語				354語	14.33
合計	1,929語	73語	115語	2,471語	

表 4
Asahi Weekly 11月27日分の CEFR-J のレベル別分布状況

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	668語	15語	18語	701語	28.91
A2	502語	10語	19語	531語	21.90
B1	496語	24語	36語	556語	22.93
B2	267語	17語	22語	306語	12.62
それ以外の語				331語	13.65
合計	1,933語	66語	95語	2,425語	

CEFR-J収録語の品詞違い語や派生語に関しても、品詞違い語は60～80語程度、派生語は90～115語程度、合計で150～190語程度であった。品詞違い語と派生語は、単独では日にちによる大きな差は見られないことがわかった。しかし、合計では、4日分を前半・後半に分けてみると、両者の最初の1日分と次の1日分との差が少し大きいことがわかった(表1～4)。

学習者向け英字新聞の使用語彙のCEFR-Jのレベル別出現頻度

A1からB2及びそれ以外の語の出現頻度を見ると、4日分の学習者向け英字新聞のいずれも、似たような状況になっている。すなわち、平均値では、A1が約67-68%、A2が約12-13%、B1が約8-9%、B2が約4%、それ以外の語が約6-7%にそれぞれなっている。このことから、4日分の各レベルの出現頻度の違いは最大でも約1%にとどまっていることがわかり、上記の分布状況と同様に記事内容の違いにかかわらず、レベル別出現頻度に違いはほとんど生じないことがわかった。また、A1からB1までで約88-90%と全体のほぼ9割を占めていることもわかった。さらに、4日とも、出現頻度が多い順に、A1、A2、B1、それ以外の語、B2という順序になっていることもわかった。このことから、記事内容の違いにかかわらず、各レベルの出現頻度の順序の違いもまったくないことがわかった(表5～8)。

表5
Asahi Weekly 11月6日分のCEFR-Jのレベル別出現頻度(%)

	最低値	最高値	平均値
A1	43.75	84.06	67.10
A2	7.42	20.83	12.74
B1	3.82	14.59	8.47
B2	1.19	10.42	4.46
それ以外の語*	3.31	17.60	7.22

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表6
Asahi Weekly 11月13日分のCEFR-Jのレベル別出現頻度(%)

	最低値	最高値	平均値
A1	55.24	79.18	67.90
A2	6.15	21.62	13.03
B1	3.65	14.68	8.58
B2	0.41	7.32	4.22
それ以外の語*	2.25	16.22	6.26

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表7
Asahi Weekly 11月20日分CEFR-Jのレベル別出現頻度(%)

	最低値	最高値	平均値
A1	51.75	80.40	68.04
A2	5.81	17.62	11.77
B1	2.39	13.54	8.60
B2	0.00	8.31	4.29
それ以外の語*	1.59	14.79	7.30

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表 8
Asahi Weekly 11月27日分 CEFR-J のレベル別出現頻度 (%)

	最低値	最高値	平均値
A1	53.00	82.42	67.26
A2	6.64	19.44	13.40
B1	3.44	15.73	8.42
B2	0.78	8.10	4.05
それ以外の語*	4.73	14.92	6.87

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

レベル別出現頻度の場合、CEFR-J 収録語以外の語の中には、品詞違い語や派生語が含まれており、表 1～4 から、その割合は約33-35%である。約6-7%であったそれ以外の語のうち、CEFR-J 収録語に関わらない語の出現頻度は、約4.2-4.8%以下と考えられる（表 5～8）。

CEFR-J 収録語以外の語の特徴

CEFR-J 収録語以外の特徴を見ると、やはり、辞書で低頻度語と定義している語が圧倒的に多い。接頭辞や接尾辞 -er の多用、正式語や新語、専門用語などと、特徴を挙げることはできるが、その多くは低頻度語である。従って、CEFR-J 収録語以外の語の大きな特徴は、低頻度語でかつ、接頭辞や接尾辞 -er の多用、正式語や新語、専門用語などと言った方が正確と言える。

しかし、CEFR-J 収録語以外の語の中には、日本で出版されている主要な英和辞典や LDOCE などの英英辞典で重要語と定義されている語も含まれている。表に挙げた重要語はいずれも高等学校の検定教科書に出現する語である（表 9）。

表 9
Asahi Weekly 4 日分の CEFR-J 収録語以外の語の特徴の例

特 徴	例
接頭辞の多用	*decarbonization; *hyperconnected; intergovernmental; <u>misinformation</u> ; *multiyear; <u>noncommittally</u> ; <u>rename</u> ; *unassume
接尾辞 -er の多用	<u>hitchhiker</u> ; <u>homeowner</u> ; *mankiller; *notetaker; <u>policymaker</u> ; <u>quitter</u> ; <u>speechwriter</u> ; <u>seafarer</u> ; <u>taker</u>
正式語	<u>decry</u> ; <u>divergent</u> ; <u>humankind</u> ; <u>indomitable</u> ; <u>remuneration</u> ; <u>seclude</u>
新語	*coronavirus
専門用語	<u>arabesque</u> ; <u>citrus</u> ; <u>domiciled</u> ; <u>finale</u> ; <u>osteoporosis</u> ; <u>quarterback</u>
英和辞典での重要語	**drift; **equivalent; **lens; **parallel; **pat; **poll; **sacrifice;
低頻度語	<u>affiliate</u> ; <u>consolidation</u> ; <u>deterioration</u> ; <u>impeachment</u> ; <u>milligram</u> ; <u>palette</u> ; <u>sandbar</u> ; <u>tribune</u>

注. 下線を引いた語は低頻度語である。

* は主要英和辞典に未収録の語である。

** は LDOCE (6th ed.) で最重要9000語のうちの上位6000語までの語である。

考 察

学習者向け英字新聞 2 紙の使用語彙の CEFR-J のレベル別分布状況の比較

まず、*Asahi Weekly* (以下、*AW*。) 4 日分の A1 から B2 及びそれ以外の語の分布状況に記事内容や日にちの違いによる差が見られないということがわかった。頁によって差が生じることはあるが、一日分を通してみると、レベル別分布状況に大きな違いがないということである。このことは *The Japan Times Alpha* (以下、*JTA*。) 5 日分にも言えることである。また、*AW* では、A1 から B1 までで約 72-74% と全体の 4 分の 3 近くを、*JTA* では約 76-78% と全体の 4 分の 3 以上を、それぞれ占めていることもわかった。その上、各レベルが占める割合の多い順を見ると、*AW* では、11 月 6 日分は A1, A2, B1, それ以外の語, B2, 11 月 13 日分は A1, A2, B1, B2, それ以外の語という順序にそれぞれなっており、11 月 20 日, 27 日分は共に、A1, B1, A2, それ以外の語, B2 という順序に、*JTA* では、12 月 3 日分を除き、A1, A2, B1, B2, それ以外の語という順序に、それぞれなっていることもわかった。このことから、2 紙に大きな差はないが、*JTA* の方が、学習者にとって読みやすいものであることがわかった。例えば CEFR-J の収録語だからといっても、B2 レベルは学習者にとってはかなり難しい語が多数を占めており、その語の派生語もまた難しい語である。従って、A1 から B1 までが占める割合が *JTA* の方が約 4% 高いことから、*AW* よりも読みやすいと言える。研究動機のところでも示した通り、語彙カバー率が 1% 違うだけで読みに大きな影響を与えることが明らかであるので、約 4% の差により、*AW* と *JTA* の両紙を読んだ場合、*AW* の方が学習者には難しいと感じるものと考えられる (表 1-4, 10; 付録 A)。

次に、CEFR-J の有効性について見てみると、CEFR-J の収録語以外の語は、*AW* の場合は約 13-14% ということで、語彙カバー率は約 86-87%、*JTA* の場合は約 10-12% ということで、語彙カバー率は約 88-90% というにそれぞれなり、2 紙ともに有効性に疑問が残る結果となった。先行研究から最低でも 95%、できれば、99% の語彙カバー率が必要なことがわかっているので、学習者向け英字新聞でのこの数値は意外なものであった。*JTA* の方が読みやすいものであることがわかるが、1 割程度が未収録語となると、仮に CEFR-J の収録語を全部覚えたとしても、学習者向け英字新聞であっても、「適切な」読みは難しいことになる (表 1-4, 10; 付録 A)。

表10
Asahi Weekly 4 日分と The Japan Times Alpha 5 日分の CEFR-J の
レベル別分布状況の平均値の比較 (%)

		最低値	最高値	平均値
<i>Asahi Weekly</i>	A1	27.62	28.91	28.25
	A2	20.88	22.96	22.07
	B1	22.01	23.03	22.55
	B2	12.62	13.88	13.53
	それ以外の語	12.60	14.33	13.61
<i>The Japan Times Alpha</i>	A1	29.23	32.32	30.88
	A2	22.70	24.00	23.38
	B1	20.20	24.11	22.28
	B2	11.94	12.77	12.37
	それ以外の語	9.77	12.03	11.10

学習者向け英字新聞 2 紙の使用語彙の CEFR-J のレベル別出現頻度の比較

まず、AW4日分のA1からB2及びそれ以外の語の出現頻度も、分布状況と同様に、記事内容や日にちの違いによる差が見られないということがわかった。頁によって差が生じることはあるが、一日分を通してみると、レベル別出現頻度に大きな違いがないということである。このことはJTA5日分にも言える。また、両紙とも、A1からB1までで約88-90%と全体のほぼ9割を占めていることもわかった。さらに、両紙とも、いずれの日も、出現頻度が多い順に、A1、A2、B1、それ以外の語、B2という順序になっていることもわかった。前節でも述べたように、CEFR-Jの収録語以外の語の大半やB2は学習者には相当難しいので、A1からB1までで全体のほぼ9割を占めているということは、学習者にとって、ある程度読みやすいものとなっていることを示していると考えられる。以上のことから、レベル別出現頻度に関しては、レベル別分布状況とは異なり、両紙による違いはほとんどないと言える(表5～8、11; 付録B)。

次に、CEFR-Jの有効性について見てみると、CEFR-Jの収録語以外の語は、AWの場合は約6-7%で、語彙カバー率は約93-94%ということに、JTAの場合は、約6-8%で、語彙カバー率は約92-94%ということにそれぞれなり、2紙ともに有効性に疑問が残る結果となった。出現頻度の場合には、CEFR-Jの収録語以外の語の中に、CEFR-Jの収録語の品詞違い語や派生語も含まれている。その割合は、AWの場合は約33-35%であるので、実際のCEFR-Jの収録語以外の語の出現頻度は約4.2-4.8%以下、JTAの場合は約32-37%であるので、実際のCEFR-Jの収録語以外の語の出現頻度は約4.0-5.3%以下とそれぞれ考えられる。以下と書いたのは、CEFR-Jの収録語の品詞違い語や派生語の出現頻度が2回以上の可能性もあるからである。この結果、CEFR-Jの収録語の語彙カバー率は、AWでは約95.2-95.8%以上、JTAでは約94.7-96.0%以上ということにそれぞれなった。前節でも述べた通り、先行研究から最低でも95%、できれば、99%の語彙カバー率が必要なことから、2紙ともに「最低限の」語彙カバー率に過ぎず、学習者向け英字新聞でも、仮にCEFR-Jの収録語を全部覚えたとしても、「適切な」読みは難しいことが、出現頻度においても明らかになった(表5～8、11～12; 付録B)。

表11
Asahi Weekly 4日分と The Japan Times Alpha 5日分の CEFR-J の
レベル別出現頻度の平均値の比較 (%)

		最低値	最高値	平均値
Asahi Weekly	A1	67.10	68.04	67.58
	A2	11.77	13.40	12.74
	B1	8.42	8.60	8.52
	B2	4.05	4.46	4.26
	それ以外の語*	6.26	7.30	6.91
The Japan Times Alpha	A1	69.22	71.18	70.16
	A2	12.49	14.09	13.07
	B1	7.95	10.39	9.04
	B2	3.50	4.43	4.03
	それ以外の語*	6.15	7.94	6.70

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表12

Asahi Weekly 4日分と The Japan Times Alpha 5日分の CEFR-J の収録語以外の語の実出現頻度の比較 (%)

		違 + 派語*	以外の全語**	違 + 派割合(%)***	実出現頻度(%)****
Asahi Weekly	11/ 6	191	553	34.54	4.73
	11/13	149	453	32.89	4.21
	11/20	188	542	34.69	4.77
	11/27	161	493	32.66	4.63
The Japan Times Alpha	11/ 5	108	294	36.73	4.57
	11/12	101	319	31.66	4.46
	11/19	114	329	34.65	4.29
	11/26	119	355	33.52	4.09
	12/ 3	127	383	33.16	5.31

注. * 違 + 派語は品詞違い語 + 派生語を表している。

** 以外の全語は品詞違い語 + 派生語を含む CEFR-J の収録語以外の全語を表している。

*** 違 + 派割合は品詞違い語 + 派生語の割合を表している。

**** 実出現頻度は CEFR-J の収録語以外の語の実出現頻度を表している。

学習者向け英字新聞2紙の CEFR-J 収録語以外の語の特徴の比較

CEFR-J の収録語以外の語の特徴は、2紙ともに同じような傾向があり、低頻度語が圧倒的に多いということである。低頻度語であることに加えて、正式語や専門用語などであることも多いと言えるので、学習者には相当難しい語が大半を占めていることがわかった。このことは、ある意味当然で、学習者向け語彙リストであるので、コーパスなど様々な題材の中で数多く出現する語を編者の知見・判断に基づいてリスト化するわけである。その編纂作業には莫大な時間を要するので、新語が含まれないことも当然と言える。(表9; 付録C)

しかし、2紙ともに、日本で出版されている主要な英和辞典や LDOCE などの英英辞典で重要語と定義されている語も収録語以外の語の中に意外と多く含まれている。中には2紙に共通して出現した語も数語あった。LDOCE の6000語には含まれないが、日本で出版されている主要な英和辞典には重要語と定義されている語も散見された。このような重要語は、学習者の今まで学習経験の中で遭遇したり、覚えたりした経験がある語も含まれている可能性は大いにある。また、新語は学習者が日常で遭遇する可能性がある。そうした語を差し引けば、学習者にとっての実際の未知語の数はある程度減る可能性が高く、上記の2つの節で述べた今回の調査から見た CEFR-J の語彙カバー率は多少上がるものと予測される。それにしてもなお、「最適な」語彙カバー率と言える99%には程遠いのである(表9, 13; 付録C)。

表13

Asahi Weekly 4日分日分と The Japan Times Alpha 5日分に見られる CEFR-J の収録語以外の語の中の英和辞典での重要語の例の比較 (%)

新聞名	例
Asahi Weekly	*drift; * <u>equivalent</u> ; *lens; *parallel; *pat; *poll; *sacrifice
The Japan Times Alpha	*cart; * <u>equivalent</u> ; *lens; *magnet; *nest; *pat; *poll; *sacrifice

注. 波線を引いた語は2紙に共通して出現した語である。

* は LDOCE (6th ed.) で最重要9000語のうちの上位6000語までの語である。

教育上の示唆

今回の調査で、2紙ともにA1からB2及びそれ以外の語の分布状況と同様に、出現頻度も記事内容や日にちの違いによる差が見られないということがわかった。頁によって差が生じることはあるが、一日分を通してみると、レベル別分布状況や出現頻度に大きな違いがないということである。従って、教材や課題として学習者に提示する場合に、日にちのことを気にすることなく提示することができることがわかった。もちろん、頁による違いがあるので、指導している学習者の実態に応じた教材や課題の選定が必要である。

また、CEFR-Jの有効性を見ると、レベル別分布状況からは、CEFR-Jの収録語の語彙カバー率は、AWでは約86-87%、JTAでは約88-90%ということが、レベル別出現頻度からは、AWでは約95.2-95.8%以上、JTAでは約94.7-96.0%ということが、それぞれわかった。JTAの方が読みやすいことがわかったが、学習者向け英字新聞という割には、語彙カバー率が2紙ともに予想よりも低いことがわかった。このことから、学習者向け英字新聞といえども、記事にある注釈以外に何の処置もしないで提示してしまうと、学習者に「適切な」読みをさせることは難しいと言えよう。

具体的な処置の例としては、八島(2022c)で述べたように、購入した学習者向け英字新聞をスキャナーで取り込み、Wordファイルとしてから、B2やCEFR-Jの収録語以外の語を、指導している学習者の実態に応じて別の語に書き換えたり、八島(2013, 2020c)で行ったように、学習者にとって難しいと判断した語の真上に日本語の訳語をつけたりする方法が考えられる。後者の方法であれば、注釈を見ることで起こる読みの中断をある程度防ぐことができると思われる。B2やCEFR-Jの収録語以外の語でA1やA2の語に書き換えが可能な語は積極的に書き換えて、両者を併用することで、読みの中断がなるべく起こらないようにすることができる。このような処置をすることで、CEFR-Jの語彙カバー率を99%にすることができれば、学習者にとって理想的な読みの教材になると思われる。学習者はそのように工夫された補助教材を多読することで速読の訓練を積むことができるであろう。

学習者の語彙知識は一人一人異なっているので、学習者全員にCEFR-Jの語彙カバー率が99%の記事を与えることは非常に難しい。しかし、学習者主体の学習の場を提供するには、学習者一人一人の実態に応じた指導をしようとする必要があるであろう。

最後に、CEFR-Jの収録語以外の語の中には、日本で出版されている主要な英和辞典やLDOCEなどの英英辞典で重要語と定義されている語が含まれていることがわかった。CEFR-Jの収録語を中心として、検定教科書やテキスト、学習者向け英字新聞などの補助教材で出現頻度がある程度高い語も、CEFR-Jの収録語以外の語であっても、英和辞典やLDOCEなどの英英辞典で確認をして、重要語と定義されている語であれば、学習者に積極的に習得するように指導すべきであろう。

今後の課題

今回、CEFR-Jの収録語と昨年とは違う学習者向け英字新聞の使用語彙との対応を調査し、昨年の研究結果との比較を行ったが、2紙ともにCEFR-Jの有効性に疑問が残る結果となった。しかし、2紙ともに対象が一月分ほどのテキストに過ぎず、その結果を受けて、CEFR-Jの有効性は疑問であると断言することはできないと言えよう。今後、少なくとも、半年分の学習者向け英字新聞2紙の調査をした上で、CEFR-Jの有効性に関する判断を下す必要があると考える。ま

た、学習者にとって望ましい語彙リストを見出すために、CEFR-Jとは別の語彙リストで今回と同様の調査をして、その結果を比較してみたい。

謝 辞

本研究は、令和2年度～令和6年度科学研究費（基盤研究（C）・課題番号：20K00881・研究代表者：八島等）の助成を受けて行われた研究の一部である。

注

本稿は、全国英語教育学会第48回香川研究大会（香川大学）及び、日本教科教育学会第49回全国大会（弘前大学）において発表した内容に加筆・訂正をしたものである。

引用・参考文献

- Carver, R. P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior*, 26(4), 413-437.
- Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689-696.
- Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Multilingual Matters.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). Macmillan.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Liu Na, & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16(1), 33-42.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59-82.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). Longman.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge University Press.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43.
- Schmitt, N., Schmitt, D., & Capham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88.
- Yashima, H. (2001). Word inference ability in Japanese EFL learners, *ARELE*, 12, 101-110.
- Yashima, H. (2002). Factors affecting guessing unknown word meanings from context, *ARELE*, 13, 151-160.
- Yashima, H. (2003). Analysis of senior high school students' ability to guessing unknown word meanings from context, *ARELE*, 14, 161-170.
- 東京外国語大学投野由紀夫研究室 (2013). 『CEFR-J Wordlist Version 1.0』.
- 東京外国語大学投野由紀夫研究室 『CEFR-J Wordlist Version1.6』. (URL: <http://www.cefr-j.org/download>).

- html#cefrj_wordlist より2020年12月8日ダウンロード)
- 投野由起夫 (2013). 『英語到達度指標 CEFR-J ガイドブック』大修館書店.
- 羽鳥博愛 (1979). 「リーディングテスト作成上の留意点」羽鳥博愛・浅野博・伊村元道・大井上滋・大友賢二・清川英男 (編) 『英語指導法ハンドブック④評価編』 (pp. 107-121) 大修館書店.
- 八島等 (1996). 「検定教科書での未知語の推測」『語学教育研究所紀要』, 10, 35-42.
- 八島等 (1999). 「未知語の割合と文脈からの意味の推測の関係及び推測に影響を与える他の要因」『語学教育研究所紀要』, 13, 31-49.
- 八島等 (2013). 「文脈からの推測のための最適な割合と語彙サイズを求めて」『日本教科教育学会第39回全国大会論文集』, 106-107.
- 八島等 (2015). 「日本人英語学習者における語彙カバー率と読解との関係」『日本教科教育学会第41回全国大会論文集』, 138-139.
- 八島等 (2016). 「これからの日本人英語学習者が習得すべき語彙の実態」『日本教科教育学会第42回全国大会論文集』, 46-47.
- 八島等 (2019a). 「CEFR-J から見た高等学校の改定版検定教科書の語彙の実態 (3)」『日本教科教育学会第45回全国大会論文集』, 204-205.
- 八島等 (2019b). 「CEFR-J から見た高等学校の改定版検定教科書の語彙の実態」『広島文教大学紀要』, 54, 65-74.
- 八島等 (2020a). 「CEFR-J の英字新聞に対する有効性の一研究」『日本教科教育学会第46回全国大会論文集』, 85-86. https://drive.google.com/drive/folders/1qUOLzVqXD5qjMDqFYjUITiq3_h8iCui?usp=sharing
- 八島等 (2020b). 「CEFR-J は英字新聞に対してどの程度有効か」第3回 JAAL in JACET 学術交流集会発表資料.
- 八島等 (2020c). 「文脈からの推測のための最適な割合と語彙サイズを求めて」『広島文教大学紀要』, 55, 1-12.
- 八島等 (2021a). 「英字新聞に対する CEFR-J の有効度に関する研究」JACET 第60回国際記念大会発表資料.
- 八島等 (2021b). 「出現頻度から見た CEFR-J の英字新聞に対する有効性」『日本教科教育学会第47回全国大会論文集』, 143-144. <https://drive.google.com/file/d/1fa7lGMdeFV8q6RETj0Wfnym8vzz1B6Y1/view?usp=sharing>
- 八島等 (2022a). 「学習者向け英字新聞に対する CEFR-J の有効度に関する研究」JACET 第61回国際記念大会発表資料.
- 八島等 (2022b). 「出現頻度から見た CEFR-J の学習者向け英字新聞に対する有効性」『日本教科教育学会第48回全国大会論文集』, 133-134. https://drive.google.com/file/d/1ETtIc-bymsDv8Qbw7Pxpjq_TaYnjG34b/view?usp=sharing
- 八島等 (2022c). 「CEFR-J の学習者向け英字新聞に対する有効性の研究」『広島文教大学紀要』, 57, 33-45.
- 八島等 (2023a). 「学習者向け英字新聞2紙における CEFR-J の有効度の比較」『全国英語教育学会第48回香川研究大会発表予稿集』, 24-25. <https://sites.google.com/shikokueigo.org/jasele2023/proceedings>
- 八島等 (2023b). 「出現頻度から見た CEFR-J の学習者向け英字新聞2紙に対する有効性の比較」『日本教科教育学会第49回全国大会論文集』, 159-160. <https://cloud.hirosaki-u.ac.jp/index.php/s/mTRZEAbx9ByKGYw/download>

付録 A : *The Japan Times Alpha* 5 日分の CEFR-J のレベル別分布状況

表 1

The Japan Times Alpha 11月 5 日分の CEFR-J のレベル別分布状況 (八島, 2022c)

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	592語	17語	6語	615語	32.32
A2	424語	10語	12語	446語	23.44
B1	389語	12語	21語	422語	22.18
B2	204語	14語	16語	234語	12.30
それ以外の語				186語	9.77
合計	1,609語	53語	55語	1,903語	

表 2

The Japan Times Alpha 11月12日分の CEFR-J のレベル別分布状況 (八島, 2022c)

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	572語	16語	11語	599語	30.55
A2	445語	8語	15語	468語	23.87
B1	409語	5語	19語	433語	22.08
B2	216語	10語	17語	243語	12.39
それ以外の語				218語	11.12
合計	1,642語	39語	62語	1,961語	

表 3

The Japan Times Alpha 11月19日分の CEFR-J のレベル別分布状況 (八島, 2022c)

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	582語	19語	13語	614語	31.88
A2	441語	7語	14語	462語	24.00
B1	366語	8語	15語	389語	20.20
B2	208語	17語	21語	246語	12.77
それ以外の語				215語	11.17
合計	1,597語	51語	63語	1,926語	

表 4

The Japan Times Alpha 11月19日分の CEFR-J のレベル別分布状況 (八島, 2022c)

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	601語	22語	7語	630語	30.42
A2	451語	11語	12語	474語	22.89
B1	446語	9語	18語	473語	22.84
B2	218語	19語	21語	258語	12.46
それ以外の語				236語	11.40
合計	1,716語	61語	58語	2,071語	

表 5

The Japan Times Alpha 12月 3 日分の CEFR-J のレベル別分布状況 (八島, 2022c)

	CEFR-J 収録語	品詞違い語	派生語	合計	割合 (%)
A1	598語	16語	8語	622語	29.23
A2	455語	12語	16語	483語	22.70
B1	471語	15語	27語	513語	24.11
B2	221語	7語	26語	254語	11.94
それ以外の語				256語	12.03
合計	1,745語	50語	77語	2,128語	

付録 B : *The Japan Times Alpha* 5 日分の CEFR-J のレベル別出現頻度

表 6

The Japan Times Alpha 11月 5 日分の CEFR-J のレベル別出現頻度 (%) (八島, 2022c)

	最低値	最高値	平均値
A1	54.29	85.95	69.66
A2	6.49	19.93	12.53
B1	3.24	13.73	7.95
B2	1.08	7.39	3.50
それ以外の語*	3.17	10.00	6.35

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表 7

The Japan Times Alpha 11月12日分の CEFR-J のレベル別出現頻度 (%) (八島, 2022c)

	最低値	最高値	平均値
A1	55.83	84.83	71.18
A2	3.93	20.95	14.09
B1	2.17	15.63	9.53
B2	0.00	12.50	4.43
それ以外の語*	1.56	10.93	6.52

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表 8

The Japan Times Alpha 11月19日分の CEFR-J のレベル別出現頻度 (%) (八島, 2022c)

	最低値	最高値	平均値
A1	54.60	85.19	69.94
A2	5.41	20.45	13.52
B1	1.35	12.38	8.60
B2	1.12	10.34	4.10
それ以外の語*	1.56	14.07	6.56

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表 9

The Japan Times Alpha 11月26日分の CEFR-J のレベル別出現頻度 (%) (八島, 2022c)

	最低値	最高値	平均値
A1	58.24	84.13	70.80
A2	7.27	16.93	12.71
B1	3.70	13.92	8.72
B2	2.11	9.41	4.11
それ以外の語*	2.91	11.11	6.15

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

表10

The Japan Times Alpha 12月 3 日分の CEFR-J のレベル別出現頻度 (%) (八島, 2022c)

	最低値	最高値	平均値
A1	44.74	79.89	69.22
A2	5.75	14.60	12.49
B1	1.35	26.97	10.39
B2	0.99	10.00	3.99
それ以外の語*	3.46	14.40	7.94

注. *それ以外の語の中には品詞違い語や派生語が含まれている。

付録 C : *The Japan Times Alpha* 5 日分の CEFR-J 収録語以外の語の特徴の例

表11

The Japan Times Alpha 5 日分の CEFR-J 収録語以外の語の特徴の例 (八島, 2022c)

特 徴	例
接頭辞の多用	<u>counterterrorism</u> ; *intraparty; <u>multimillion</u> ; *nonconform; <u>restart</u> ; <u>transcontinental</u> ; <u>unintended</u>
接尾辞 -er の多用	<u>automaker</u> ; *bombmaker; <u>dispatcher</u> ; <u>theatergoer</u> ; <u>whistleblower</u>
正式語	<u>confection</u> ; <u>designation</u> ; <u>expiry</u> ; <u>incapacitate</u> ; <u>inalienable</u> ; <u>unquestionably</u>
新語	*coronavirus; *telework
専門用語	<u>anesthesia</u> ; <u>bumblebee</u> ; <u>helium</u> ; <u>lotus</u> ; <u>mastodon</u> ; <u>postpartum</u> ; <u>therapeutics</u>
英和辞典での重要語	**cart; **equivalent; **lens; **magnet; **nest; **pat; **poll; **sacrifice
低頻度語	<u>assassination</u> ; <u>baguette</u> ; <u>chocaholic</u> ; <u>excavation</u> ; <u>lampshade</u> ; <u>ratchet</u> ; <u>skeptic</u> ; <u>temblor</u>

注. 下線を引いた語は低頻度語である。

* は主要英和辞典に未収録の語である。

** は LDOCE (6th ed.) で最重要9000語のうちの上位6000語までの語である。

—2023年 9 月25日 受理—