

【原著】

学習者向け英字新聞に対する 新 JACET8000 と CEFR-J の有効性の比較

八 島 等

A Comparison of the Effectiveness of the New JACET 8000
and the CEFR-J Wordlist on Newspapers for Learners of English

Hitoshi Yashima

一昨年、CEFR-Jで示されているA1からB2までの語彙が学習者向け英字新聞でどの程度用いられているのかを調査した。その結果、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%であったが、CEFR-Jの語の出現頻度は一日分平均で約92-94%であった。昨年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%となったが、出現頻度は約93-94%であった。2紙ともに、レベル別分布状況ではCEFR-Jの有効性に疑問が残ったが、出現頻度は「最低限の」語彙カバー率に近いものであった。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。一昨年の調査では約32-37%、昨年の調査では約33-35%それぞれ含まれていたため、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た。今年には新JACET8000を用いて昨年と同じ学習者向け英字新聞を調査すると、1000-3000が約70-71%、4000-5000が約12%、6000-8000が約8-9%、それ以外が約9-10%であったが、出現頻度は約96-97%であった。実際の出現頻度は約98%以上であったことから、ほぼ「最適な」語彙カバー率と言える。

キーワード：新JACET8000, CEFR-J, 学習者向け英字新聞, 有効性, 比較

研究動機

11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、9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、一昨年、学習者向け英字新聞*The Japan Times Alpha*の一个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)。

さらに、昨年、学習者向け英字新聞*Asahi Weekly*の一个月分を研究対象として上記の学習者向け英字新聞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。学習者にとって、一昨年度の学習者向け英字新聞の方が読みやすいということがわかり、CEFR-Jの有効性に疑問が残った。しかし、CEFR-Jの語の出現頻度は、一日分平均で、約93-94%であり、昨年同様「最低限の」語彙カバー率に近いものであった。その上、すべてのそれ以外の語のうちの33-35%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4.2-4.8%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, 2023a, b)。

研究目的

本研究の主な目的は、新JACET8000に収められている語がどの程度、昨年と同じ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、昨年使用した学習者向け英字新聞における新JACET8000の収録語のレベル別分布状況と新JACET8000の収録語以外の語の出現率を調査して、新JACET8000の有効性と新JACET8000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、大学英語教育学会基本語改訂特別委員会(編)の『大学英語教育学会基本語リスト新JACET8000』(以下、新JACET8000)、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』、及び購入した*Asahi Weekly*の2022年11月6、13、20、27日版である。

新JACET8000にはその名の通り、8000語を収録しており、その8000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、新JACET8000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、昨年のCEFR-Jを用いた調査に準じた。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。

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。その上で再度新JACET8000と対照し、1語1語レベルの確認をして、品詞違い語と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。新JACET8000の1000-8000語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、新JACET8000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、新JACET8000の収録語以外の語の中に、新JACET8000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、新JACET8000の有効性を検証して、昨年調査結果と比較した。さらに、新JACET8000の未収録語の中に重要語¹がどの程度含まれているのかも調査した。

結 果

学習者向け英字新聞の使用語彙の新JACET 8000のレベル別分布状況

1000語から8000語（以下「語」を省略）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、1000-3000が約70-71%、4000-5000が約12%、6000-8000が約8-9%、それ以外が約9-10%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1%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、レベル別分布状況に違いはほとんど生じないことがわかった。また、1000-5000で約82-83%と全体の5分の4以上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1回おきに分かれ、11月6、20日分は共に、1000、2000、3000、4000-5000、それ以外の語、6000-8000という順序にそれぞれなっているが、11月13、27日分は共に、1000、2000、4000-5000、3000、それ以外の語、6000-8000という順序にそれぞれ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1%未満であることから、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日分の新JACET8000のレベル別分布状況

	新JACET8000収録語	品詞違い語	派生語	合計	割合 (%)
1000語	788語	144語	10語	942語	35.68
2000語	493語	43語	8語	544語	21.05
3000語	295語	25語	12語	332語	12.85
4000-5000語	290語	15語	9語	314語	12.15
6000-8000語	194語	9語	14語	217語	8.40
それ以外の語				255語	9.87
合計	2060語	236語	53語	2584語	

¹ 本論文では、重要語は以下の①と②を共に満たすものとする。①最低でも、Genius 6及びWisdom 4のいずれの辞書においても、中学または高校の必修語となっている語であること。②LDOCE (6th ed.) の上位6000語までの語であること。

表 2

Asahi Weekly 11月13日分の新JACET8000のレベル別分布状

	新JACET8000収録語	品詞違い語	派生語	合計	割合 (%)
1000語	762語	155語	4語	921語	37.75
2000語	492語	45語	1語	538語	22.05
3000語	265語	11語	6語	282語	11.56
4000-5000語	266語	9語	12語	287語	11.76
6000-8000語	189語	1語	8語	198語	8.11
それ以外の語				214語	8.77
合計	1974語	221語	31語	2440語	

表 3

Asahi Weekly 11月20日分の新JACET8000のレベル別分布状況

	新JACET8000収録語	品詞違い語	派生語	合計	割合 (%)
1000語	740語	157語	6語	903語	36.81
2000語	454語	50語	6語	510語	20.79
3000語	270語	19語	10語	299語	12.19
4000-5000語	268語	16語	7語	291語	11.83
6000-8000語	177語	4語	12語	193語	7.81
それ以外の語				257語	10.48
合計	1909語	246語	41語	2453語	

表 4

Asahi Weekly 11月27日分の新JACET8000のレベル別分布状況

	新JACET8000収録語	品詞違い語	派生語	合計	割合 (%)
1000語	750語	159語	18語	927語	37.87
2000語	455語	51語	10語	516語	21.08
3000語	252語	15語	4語	271語	11.07
4000-5000語	264語	14語	15語	293語	11.97
6000-8000語	197語	4語	12語	213語	8.70
それ以外の語				227語	9.27
合計	1918語	243語	59語	2448語	

新JACET8000収録語の品詞違い語や派生語に関しても、品詞違い語は220～250語程度、派生語は30～60語程度、合計で250～300語程度であった。品詞違い語と派生語は、単独では日ごとの大きな差は見られないことがわかった。しかし、合計では、11月13日分と11月27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新JACET8000のレベル別出現頻度

1000-8000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、1000-3000が約89-91%、4000-5000が約4%、6000-8000が約2-3%、それ以外の語が約3-4%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2%であることがわかり、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、1000-5000で約93-95%と全体の9割5分近くを占めていることもわかった。さらに、11月13日を除いて、出現頻度が多い順に、1000、2000、3000、それ以外の語、4000-5000、6000-8000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ほぼないことがわかった(表5～8)。

表 5

Asahi Weekly 11月6日分の新JACET8000のレベル別出現頻度 (%)

	最低値	最高値	平均値
1000語	56.25	86.32	75.10
2000語	4.27	14.58	9.72
3000語	.92	10.42	4.43
4000-5000語	.85	10.41	4.14
6000-8000語	.54	6.53	2.42
それ以外の語*	.54	9.30	4.33

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日分の新JACET8000のレベル別出現頻度 (%)

	最低値	最高値	平均値
1000語	59.46	84.43	76.74
2000語	5.22	19.48	10.16
3000語	1.16	10.81	4.40
4000-5000語	1.16	8.11	3.53
6000-8000語	.40	4.02	2.23
それ以外の語*	.58	8.11	3.30

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 7

Asahi Weekly 11月20日分の新JACET8000のレベル別出現頻度 (%)

	最低値	最高値	平均値
1000語	57.47	88.10	75.43
2000語	4.76	16.10	9.58
3000語	.00	8.74	4.56
4000-5000語	.00	8.33	4.15
6000-8000語	.41	5.95	2.48
それ以外の語*	1.58	10.34	4.21

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日分の新JACET8000のレベル別出現頻度 (%)

	最低値	最高値	平均値
1000語	64.00	85.60	74.99
2000語	4.77	17.66	10.08
3000語	.77	8.28	4.47
4000-5000語	.99	12.00	4.03
6000-8000語	.35	7.14	2.72
それ以外の語*	2.00	10.09	4.21

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、新JACET8000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53-57%である。約3-4%であったそれ以外の語のうち、新JACET8000収録語に関わらない語の出現頻度は、約1.5-2.0%以下と考えられる(表5～8)。

新JACET8000収録語以外の語の特徴

新JACET8000収録語以外の特徴を見ると、CEFR-Jの場合と同様にやはり、辞書で低頻度語と定義している語が圧倒的に多い。また、主要英和辞典に未収録の語も散見された。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かなり含まれていた。そこで、新JACET8000収録語以外の語の中の重要語を拾い上げるとわずかに3語であった。(表9)。

表9
Asahi Weekly 4日分の新JACET8000収録語以外の語の特徴

特 徴	例
接頭辞の多用	*decarbonization; *hyperconnected; *multiyear; <u>noncommittally</u> ; rename; <u>undefined</u> etc.
接尾辞-erの多用	*decipherer; filmmaker; <u>hitchhiker</u> ; *mankiller; *notetaker; <u>quitter</u> ; speechwriter; <u>seafarer</u> ; taker etc.
正式語	decry; divergent; humankind; <u>indomitable</u> ; <u>remuneration</u> ; seclude etc.
新語	*coronavirus
専門用語	<u>arabesque</u> ; citrus; <u>domiciled</u> ; mutation; osteoporosis; truss etc.
重要語	***cellphone; ***o'clock; **yawn
低頻度語	<u>abbreviation</u> ; <u>confectionery</u> ; <u>decimation</u> ; <u>ineptitude</u> ; <u>prefectural</u> ; <u>sandbar</u> ; <u>tweet</u> ; <u>vandalize</u> etc.

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。

***はLDOCE (6th ed.) で最重要9000語のうちの上位3000語までの語である。

考 察

学習者向け英字新聞の使用語彙の新JACET8000とCEFR-Jのレベル別分布状況の比較

まず、新JACET8000の1000-8000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CEFR-JのA1-B2及びそれ以外の語にも言えることであった。また、新JACET8000の1000-5000までで約82-83%と全体の5分の4以上を占めていることもわかった。CEFR-Jの場合は、A1-B1まで(5022項目)で約72-74%と全体の4分の3近くしか占めていなかった。例え新JACET8000やCEFR-Jの収録語だからといっても、6000-8000や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いる。従って、新JACET8000の1000-5000の語彙カバー率の方がCEFR-JのA1-B2よりも約10%も高いことということは、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、新JACET8000の方が、学習者にとって有効性が高いことがわかった(表1~4, 10; 付録A)。

次に、新JACET8000の各レベルが占める割合の多い順を見ると、11月6、20日分は共に、1000、2000、3000、4000-5000、それ以外の語、6000-8000という順序にそれぞれなっているが、11月13、27日分は共に、1000、2000、4000-5000、3000、それ以外の語、6000-8000という順序にそれぞれなっていることがわかった。CEFR-Jの場合は、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いるが、

11月20日, 27日分は共に, A1, B1, A2, それ以外の語, B2という順序にそれぞれなっていることもわかった。このことから, 11月6, 27日分は両方の語彙リストに共通の傾向が見られ, 11月13, 20日分は両方の語彙リストに異なる傾向が見られた。しかし, 両方の語彙リストの収録語の各レベルが占める割合の多い順に大きな差はなかったので, 各レベルが占める割合の多い順からは新JACET8000の方が, 学習者にとって有効性が高いとは言い切れないことがわかった(表1~4, 10; 付録A)。

表10
Asahi Weekly 4日分の新JACET8000と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
新JACET8000	1000語	35.68	37.87	37.03
	2000語	20.79	22.05	21.24
	3000語	11.07	12.85	11.92
	4000-5000語	11.76	12.15	11.90
	6000-8000語	7.87	8.70	8.27
	それ以外の語	8.77	10.48	9.60
CEFR-J	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

学習者向け英字新聞の使用語彙の新JACET8000とCEFR-Jのレベル別出現頻度の比較

まず, 新JACET8000の1000-8000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CEFR-JのA1-B2及びそれ以外の語にも言えることであった。また, 新JACET8000の1000-5000までで約93-95%と全体の9割5分近くを占めており, この割合は「最低限の」語彙カバー率に近いものであった。CEFR-Jの場合は, A1-B1まで(5022項目)で約88-90%と全体のほぼ9割を占めていた。前節でも述べた通り, 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, 新JACET8000の1000-5000の語彙カバー率の方がCEFR-JのA1-B2よりも約5%も高いということは, 新JACET8000の方が, 学習者にとって有効性が高いことがわかった。それに対して, 新JACET8000の各レベルが占める割合の多い順を見ると, いずれの日も, 1000, 2000, 3000, それ以外の語, 4000-5000, 6000-8000という順序になっていることもわかった。CEFR-Jの場合は, いずれの日も, 多い順に, A1, A2, B1, それ以外の語, B2という順序になっていた。以上のことから, 両方の語彙リストの収録語の各レベルが占める割合の多い順に違いはまったくなかったので, 各レベルが占める割合の多い順からは, 新JACET8000の方が学習者にとって有効性が高いとは言えないことがわかった(表5~8, 11; 付録B)。

次に, 両方の語彙リストの収録語以外の語の実出現頻度について見てみる。両方の語彙リストとも, 出現頻度の場合には, 収録語以外の語の中に, 収録語の品詞違い語や派生語も含まれている。新JACET8000の収録語以外の語の出現頻度は約3-4%であったが, その中に, 品詞違い語や派生語が約53-57%含まれていたもので, 実出現頻度は約1.5-2.0%以下と考えられる。以下と述べたのは, 新JACET8000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。従って, 語彙カバー率は約98.0-98.5%以上ということになり, ほぼ「最適」語彙カ

パー率と言える。一方で、CEFR-Jの収録語以外の語の出現頻度は約6-7%であったが、その中に、品詞違い語や派生語が約33-35%含まれていたため、実出現頻度は約4.2-4.8%以下であった。この結果、語彙カバー率は、約95.2-95.8%以上ということになり、「最低限の」語彙カバー率と言えた。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、学習者向け英字新聞であれば、新JACET8000の1000-5000の収録語を全部覚えれば、「適切な」読みがかなりできることが期待されることが出現頻度から明らかになった。(表5～8, 11～12; 付録B)。

表11
Asahi Weekly 4日分の新JACET8000と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
新JACET8000	1000語	74.99	76.76	75.57
	2000語	9.58	10.16	9.89
	3000語	4.40	4.56	4.47
	4000-5000語	3.53	4.15	3.96
	6000-8000語	2.23	2.72	2.46
	それ以外の語	3.30	4.21	4.01
CEFR-J	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

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日分の新JACET8000と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 (%)****
新JACET8000	11/6	284	544	53.13	2.03
	11/13	252	466	54.08	1.52
	11/20	287	544	52.76	1.99
	11/27	302	531	56.87	1.82
CEFR-J	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

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+派生語を含む新JACET8000またはCEFR-Jの収録語以外のすべての語を表している。

***違+派割合は品詞違い語+派生語の割合を表している。

****実出現頻度は新JACET8000またはCEFR-Jの収録語以外の語の実際の出現頻度を表している。

学習者向け英字新聞に見られる新JACET8000とCEFR-Jの収録語以外の語の中の重要語の比較

新JACET8000の収録語以外の語の特徴は、昨年のCEFR-Jの結果と同じ傾向があり、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)

そこで、両方の語彙リストにおいて顕著な違いが見られた重要語に焦点を当てると、新JACET 8000の収録語以外の語の中に重要語はわずかに3語しか含まれていなかった。その中の2語はCEFR-Jの収録語であったので、この結果だけ見ると、CEFR-Jの方が学習者にとって有効性が高いように思われる。しかし、CEFR-Jの収録語以外の語の中には重要語が14語も含まれており、新JACET8000の方が学習者にとって有効性が高いことは明らかである。さらに、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に関しても、重要語と同様に、LDOCE の上位3000語を除いて、新JACET8000の方が圧倒的に収録されていることがわかった。とりわけ、ジーニアスの高校必修語に関しては、新JACET8000にはCEFR-Jの10倍以上の収録語があった。このことから、新JACET8000の方が学習者にとって有効性が高いことは明らかと言える（表13～14; 付録C）。

表13

Asahi Weekly 4日分に見られる新JACET8000とCEFR-Jの収録語以外の語の中の重要語の比較

新聞名	例
新JACET8000	*** <u>cellphone</u> ; *** <u>o'clock</u> ; <u>yawn</u> (3語)
CEFR-J	*** <u>cellphone</u> ; <u>communist</u> ; <u>drift</u> ; <u>equivalent</u> ; *** <u>etc.</u> ; <u>lens</u> ; <u>parallel</u> ; <u>pat</u> ; <u>poll</u> ; <u>sacrifice</u> ; <u>saddle</u> ; <u>saint</u> ; <u>subsequent</u> ; *** <u>zero</u> (14語)

注. ***はLDOCE (6th ed.) で最重要9000語のうちの上位3000語までの語である。
点線を引いた語はWebsterの3000語である。

表14

Asahi Weekly 4日分に見られる新JACET8000とCEFR-Jの収録語以外の語の中の「重要語」の比較

辞典名 (レベル)	新JACET8000	CEFR-J	共通
LDOCE (3000語)	<u>euro</u> ; <u>litter</u> ; <u>o'clock</u> (3語)	<u>etc.</u> ; <u>zero</u> (2語)	<u>cellphone</u> ; <u>mummy</u> (2語)
ジーニアス (高校必修語)	<u>deceive</u> ; <u>dime</u> ; <u>overtake</u> ; <u>pence</u> ; <u>yawn</u> (5語)	<u>clay</u> ; <u>grind</u> ; <u>mode</u> etc. (90語)	<u>antennae</u> ; <u>cellphone</u> ; <u>erase</u> ; <u>moth</u> (4語)
ウィズダム (高校必修語)	<u>ache</u> ; <u>euro</u> ; <u>litter</u> ; <u>overtake</u> ; <u>yawn</u> (5語)	<u>bundle</u> ; <u>chamber</u> ; <u>lord</u> ; etc. (30語)	<u>cellphone</u> ; <u>grumble</u> (2語)

注. 点線を引いた語はWebsterの3000語である。

教育上の示唆

今回の調査で、昨年のCEFR-Jの場合と同様に、新JACET8000の1000-8000及びそれ以外の語の分布状況や出現頻度に、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、新JACET8000の有効性を見ると、新JACET8000の収録語の語彙カバー率は、レベル別分布状況からは約90-91%ということが、レベル別出現頻度からは約98.0-98.5%以上ということが、それぞれわかった。昨年の調査では、CEFR-Jの収録語の語彙カバー率は、レベル別分布状

況では約86-87%、レベル別出現頻度では約95.2-95.8%以上であったので、新JACET8000の方が学習者にとって有効性が高いことがわかった。レベル別出現頻度で約98.0-98.5%以上のカバー率ということはほぼ「最適な」語彙カバー率と言える。さらに、新JACET8000の1000-5000までで約93-95%と全体の9割5分近くを占めており、この割合は「最低限の」語彙カバー率に近いものであった。新JACET8000の収録語以外の語の約半数や6000-8000は学習者には相当難しいので、学習者にまずは新JACETの1000-5000までを言語活動を通して習得させて、使えるようになった5000語の上に、残りの語彙を習得させることが学習者にとって望ましいと考えられる。

1000-5000を習得させるまでは、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。例えば、1000-3000の習得では、語彙カバー率は約90%程度であるので、「適切な」読みはできない。そこで、具体的な処置の例としては、八島(2022c, 2023c)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、6000-8000や新JACET8000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。6000-8000や新JACET8000の収録語以外の語を1000-3000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。習得が進んでくれば4000-5000に書き換えることで学習者が挑戦する意欲が高まることが期待される。このような処置をすることで、新JACET8000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、学習者全員に新JACET8000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、新JACET8000の収録語以外の語の中には、CEFR-Jほどではないが、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。新JACET8000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、新JACET8000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、昨年研究対象とした学習者向け英字新聞の使用語彙と新JACET8000との対応を調査し、昨年のCEFR-Jを用いた研究結果との比較を行ったが、対象が一月分のテキストに過ぎず、その結果を受けて、両方の語彙リストの有効性を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、両方の語彙リスト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、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49回福岡研究大会（福岡工業大学）、JACET第63回国際記念大会（愛知大学）及び、日本教科教育学会第50回全国大会（筑波大学）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 MARRERS.
- 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. *RELIC 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.
- 大学英語教育学会基本語改訂特別委員会（編）（2016）. 『大学英語教育学会基本語リスト新JACET8000』 桐原書店.
- 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 Version 1.6』(URL: http://www.cefr-j.org/download.html#cefrj_wordlistより2020年12月8日ダウンロード)
- 投野由紀夫(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/1qUOLzvqXD5jqcMDqFYjUITiq3_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>
- 八島等(2023c).「学習者向け英字新聞2紙におけるCEFR-Jの有効性の比較」『広島文教大学紀要』, 58, 41-55.
- 八島等(2024a).「学習者向け英字新聞に出現する重要語に対する語彙リストによる対応の違い-CEFR-Jと新JACET8000の比較-」『全国英語教育学会第49回福岡研究大会発表予稿集』, 354-355. https://sec.tobutoptours.co.jp/2024/jasele49_yoyaku/Proceedings.pdf
- 八島等(2024b).「学習者向け英字新聞に対する新JACET8000とCEFR-Jの有効度の比較」JACET第63回国際記念大会発表資料.
- 八島等(2024c).「出現頻度から見た学習者向け英字新聞に対する新JACET8000とCEFR-Jの有効性の比較」『日本教科教育学会第50回全国大会論文集』, 223-224. <https://pub.conf.it.atlas.jp/ja/event/jcrda50>

付録A : *Asahi Weekly* 4日分のCEFR-Jのレベル別分布状況

付表 1

Asahi Weekly 11月6日分のCEFR-Jのレベル別分布状況 (八島, 2023c)

	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	2062語	80語	111語	2614語	

付表 2

Asahi Weekly 11月13日分のCEFR-Jのレベル別分布状況 (八島, 2023c)

	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	1960語	60語	89語	2413語	

付表 3

Asahi Weekly 11月20日分のCEFR-Jのレベル別分布状況 (八島, 2023c)

	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	1929語	73語	115語	2471語	

付表 4

Asahi Weekly 11月27日分のCEFR-Jのレベル別分布状況 (八島, 2023c)

	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	1933語	66語	95語	2425語	

付録B : *Asahi Weekly* 4 日分のCEFR-Jのレベル別出現頻度

付表 5

Asahi Weekly 11月 6 日分のCEFR-Jのレベル別出現頻度 (%) (八島, 2023c)

	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。

付表 6

Asahi Weekly 11月13日分のCEFR-Jのレベル別出現頻度 (%) (八島, 2023c)

	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。

付表 7

Asahi Weekly 11月20日分のCEFR-Jのレベル別出現頻度 (%) (八島, 2023c)

	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。

付表 8

Asahi Weekly 11月27日分のCEFR-Jのレベル別出現頻度 (%) (八島, 2023c)

	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。

付録C : *Asahi Weekly* 4日分のCEFR-J収録語以外の語の特徴の例

付表 9

Asahi Weekly 4日分のCEFR-J収録語以外の語の特徴の例 (八島, 2023cを一部訂正)

特 徴	例
接頭辞の多用	*decarbonization; *hyperconnected; intergovernmental; <u>misinformation</u> ; *multiyear; <u>noncommittally</u> ; rename; <u>unassuming</u>
接尾辞-erの多用	<u>hitchhiker</u> ; homeowner; *mankiller; *notetaker; <u>policymaker</u> ; <u>quitter</u> ; speechwriter; <u>seafarer</u> ; taker
正式語	decry; divergent; humankind; <u>indomitable</u> ; <u>remuneration</u> ; seclude
新語	*coronavirus
専門用語	<u>arabesque</u> ; citrus; <u>domiciled</u> ; finale; osteoporosis; quarterback
英和辞典での「重要語」	**drift; **equivalent; **lens; **parallel; **pat; **poll; **sacrifice
低頻度語	<u>abbreviation</u> ; <u>confectionery</u> ; <u>decimation</u> ; <u>ineptitude</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。

—2024年 9月24日 受理—