

『児童・生徒作文コーパス』を用いた漢字使用能力の推定

宮城 信(富山大学人間発達科学部)[†]

今田 水穂(文部科学省初等中等教育局)

Estimation of the Ability to Use Kanji Using “A Written Composition Corpus of Japanese Elementary and Junior High School Students”

Shin Miyagi (University of Toyama)

Mizuho Imada (Ministry of Education, Culture, Sports, Science and Technology)

要 旨

本発表では、構築中の『児童・生徒作文コーパス』を用いて、児童・生徒の作文における漢字の使用実態と漢字使用能力の推定を試みる。このコーパスは小学校1年生から中学校3年生までの児童・生徒の作文を収集、電子化した100万語規模のコーパスで、児童・生徒の言語使用実態を縦断的に調査することができる。このコーパスを用いて、漢字種別(学年別配当漢字、常用漢字、常用外漢字)、品詞、語種などの観点から、児童・生徒の学齢別の漢字使用実態を調査する。また、『現代日本語書き言葉均衡コーパス』(BCCWJ)や、大学生の書いた作文と対照することにより、作文の文体的特徴や、大学生の漢字使用実態を到達目標とした児童・生徒の漢字使用能力の発達過程の分析を行う。

1. はじめに —本研究の目的—

これまでの諸研究では、それぞれアプローチは異なるが、学習過程で習得に適した語彙はどれかという観点から考察が進められている。河内(2015)や田中(2011)は、「国語政策・国語教育のよりどころとなるような重要語彙リストを作成する」(田中、p.86)という文言に見られるように、子ども達が優先的に学習すべき語彙の選定を念頭として、日常生活における重要語彙を検討している。また、鈴木(2011)では、中等教育課程で生徒達の語の使用を調査し、それらの重要性を検討している。

もちろん国語教育の現場において、学習に適した語彙の選定は重要である。一方で、漢字使用能力が大きく伸びていく小学校中学年から中学校にかけての発達過程の調査は、管見の限りほぼ無い。そこで本研究では、どのような語句を学習すべきかという視点ではなく、子ども達が作文する際にどのような漢字をどのように使用しているのか、すなわち児童・生徒の漢字使用能力に注目する。本研究で想定する漢字使用能力は以下のようなものである。

- ・ 児童・生徒が作文する際に、語の表記にどの程度漢字を使用するのか、または選択可能であるのか、という表現に関わる能力

この能力の推定のために、本発表では小学1年から中学3年生までの児童・生徒の作文を収集、電子化したコーパスを用いて、次のような言語使用実態を調査する。

[†] miyagi@edu.u-toyama.ac.jp

- ・ 教育漢字や常用漢字の使用頻度は、学齢の進行に対してどのように変化していくのか。
- ・ 品詞別、語種別の漢字使用率は、学齢の進行に対してどのように変化していくのか。
- ・ 上位頻度語の漢字使用率は、学齢の進行に対してどのように変化していくのか。
- ・ 漢字使用の観点から見た作文の文体的特徴はどのようなものか。
- ・ 作文時に使用される漢字の種類と頻度は、どの時期にどの程度の水準で飽和するのか。

児童・生徒が自分の力だけで作成した文章は、彼らの漢字使用の実態を調べるために適した資料であるが、作文での漢字使用が、ただちに漢字使用能力を意味するわけではない。例えば、漢語は漢字書きが普通だが、和語は仮名書きでも違和感が少ないので、漢字を知っていても使わない、ということがあり得る。そこで、文字単位の漢字使用頻度だけではなく、品詞別、語種別、単語別の漢字使用率も併せて調査する。

また、作文における漢字使用の実態は、ただちに児童・生徒に求められる漢字能力を推定する資料とはならない。作文は書き言葉の多様な言語使用域(レジスタ)の一つに過ぎず、また児童・生徒の漢字使用能力の最終的な到達目標は学校教育の過程の先に位置する。そこで、『現代日本語書き言葉均衡コーパス』(BCCWJ)や大学生の作文と比較することにより、レジスタ横断的な観点から見た作文の文体的特徴や、学齢縦断的観点から見た漢字使用実態の飽和過程を分析する。

これらの調査により、児童・生徒の書く文章で要求される学齢別の標準的な漢字使用能力の範囲を推定する。また、それによって、現場での漢字学習や指導における重点化の判断や重要語の選出への示唆を与える。

2. 『児童・生徒作文コーパス』の概要

2. 1 調査の概要

国立大学附属小・中学校を調査協力校として、4校(小学校2校、中学校2校)9学年(小学1年～中学3年)の全児童・生徒に作文課題を課し(作成時間は小学校40分、中学校45分)、収集して電子化した。作文は「ゆめ」などのテーマ(タイトル)のみを提示し、教員は一切の事前指導を行わない。調査は、2014年度に2回実施した。

第1回調査:「ゆめ」、2014年7月実施

第2回調査:「ぼくの／わたしのがんばったこと」、同12月実施

平文テキストへの電子化は以下の指針に従って実施した。

○電子化の指針

- ・ できるだけ、正確に紙面を再現するよう心がける。
- ・ 段落初めの一字下げや空欄(意味不明なものも含めて)も正確に記録する。
- ・ 誤字・脱字、文字種の違いにも注意して、正確に記録する。
- ・ 入力後に入力者以外の者が原本と照合し、入力ミスを修正する。
- ・ 個人情報にかかわる部分(個人が特定される可能性のある語句や学校名、氏名・渾名など)は、当該部分を“*”で置き換える。
- ・ 1作文1ファイルで記録し、整理番号を付す。(整理番号から、課題・学年・クラス・性別などが判別できるようにする)

2. 2 データの概要

2015年7月現在の時点で電子化が完了しているテキストについて、構文解析を実施した結果を以下に示す。解析には、CaboCha 0.69、UniDic 2.1.2を使用した。

【表1】データの概要

| 課題 | 作文数 | 文数 | 文節数 | 短単位数 | 文字数 |
|---------|-------|--------|---------|-----------|-----------|
| ゆめ | 1,818 | 27,006 | 217,376 | 589,772 | 924,604 |
| がんばったこと | 1,599 | 27,829 | 208,208 | 580,151 | 922,914 |
| 計 | 3,417 | 54,835 | 425,584 | 1,169,923 | 1,847,518 |

データの作文数が異なるため、1作文あたりの数も集計し、以下の結果を得た。

【表2】データの概要 (1作文あたり)

| 課題 | 文数 | 文節数 | 短単位数 | 文字数 |
|---------|------|-------|-------|-------|
| ゆめ | 14.9 | 119.6 | 324.4 | 508.6 |
| がんばったこと | 17.4 | 130.2 | 362.8 | 577.2 |
| 平均 | 16.2 | 124.9 | 343.6 | 542.9 |

1作文あたりに換算すると、文数、文節数など今回調査した全ての項目において「がんばったこと」の方が数値が大きい。「ゆめ」は7月、「がんばったこと」は12月時点での調査である。数値の違いは、課題の違いによる可能性と、調査時期の違いによる可能性があるが、ここでは諸元の提示に留める。なお1作文あたりの平均的な分量は400字詰め原稿用紙1.3枚程度である。

3. 『児童・生徒作文コーパス』における漢字の使用実態

3. 1 学年別の漢字の使用頻度

学年別の教育漢字(小学校6年生までの学齢別配当漢字)、常用漢字(配当外)、常用外漢字の使用実態を以下に示す。数値は2課題の平均で、以下の調査も同様である。

【表3】学年別の漢字使用頻度(1万字あたり)

| | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 |
|------|-----|-----|------|------|------|------|------|------|------|
| 配当1年 | 113 | 361 | 438 | 420 | 477 | 534 | 581 | 554 | 588 |
| 配当2年 | 27 | 238 | 459 | 504 | 579 | 640 | 692 | 688 | 702 |
| 配当3年 | 9 | 10 | 210 | 290 | 365 | 414 | 455 | 415 | 472 |
| 配当4年 | 6 | 3 | 25 | 117 | 182 | 232 | 272 | 264 | 292 |
| 配当5年 | 3 | 2 | 5 | 15 | 77 | 145 | 190 | 203 | 222 |
| 配当6年 | 1 | 1 | 28 | 44 | 64 | 121 | 142 | 151 | 146 |
| 配当外 | 4 | 3 | 6 | 14 | 35 | 81 | 136 | 171 | 198 |
| 常用外 | 1 | 0 | 0 | 1 | 1 | 2 | 5 | 7 | 8 |
| 計 | 165 | 618 | 1171 | 1405 | 1780 | 2169 | 2474 | 2454 | 2627 |

表3から、小学校1年生の時点では、あまり漢字を用いず文章を書いているが、学齢が進むにつれて漢字の使用頻度が上がっていく様子が分かる。特に、小学校の低～中学年の

間は一定のペースで漢字の使用量が上昇する。教育漢字に関しては、小学 6 年次頃には、ほぼ変化しなくなり、一定程度定着したと見ることができそうである。

学習漢字を用いた漢字書きの発達をさらに詳細に見るため、中学 3 年の漢字使用状況を基準として学齢別に学習漢字の定着状況を以下に示す(中 3 の使用頻度を分母として百分率を計算した²。70%を超える学齢に下線を引いた)。

【表 4】学年別の漢字使用頻度(中 3 を 100%として) [単位: %]

| 配当学年 | 小 1 | 小 2 | 小 3 | 小 4 | 小 5 | 小 6 | 中 1 | 中 2 | 中 3 |
|--------|------|------|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| 配当 1 年 | 17.6 | 61.0 | <u>74.5</u> | 71.6 | 81.1 | 90.4 | 98.9 | 94.5 | 100.0 |
| 配当 2 年 | 3.9 | 33.8 | 65.3 | <u>71.8</u> | 82.4 | 91.2 | 98.6 | 98.0 | 100.0 |
| 配当 3 年 | 2.0 | 2.3 | 44.0 | 61.9 | <u>78.7</u> | 88.8 | 97.3 | 88.9 | 100.0 |
| 配当 4 年 | 2.1 | 1.0 | 8.4 | 39.5 | 62.8 | <u>79.9</u> | 93.9 | 91.1 | 100.0 |
| 配当 5 年 | 1.3 | 0.6 | 2.2 | 7.2 | 34.6 | 64.8 | <u>86.4</u> | 92.3 | 100.0 |
| 配当 6 年 | 0.9 | 0.6 | 19.7 | 30.0 | 44.2 | <u>83.9</u> | 98.2 | 104.0 | 100.0 |
| 常用 | 2.3 | 1.7 | 2.7 | 7.1 | 17.5 | 39.9 | 66.9 | <u>85.5</u> | 100.0 |
| 常用外 | 5.2 | 2.8 | 5.6 | 7.1 | 9.4 | 30.2 | 69.5 | <u>96.5</u> | 100.0 |
| 計 | 6.3 | 23.5 | 44.6 | 53.5 | 67.8 | <u>82.5</u> | 94.2 | 93.4 | 100.0 |

教育漢字の使用頻度は、配当学年から 2 年程度で大学生の使用頻度の 70%に達する (例えば、小 1 配当の漢字が 70%を超えるのは、小 3 年次である)。中 1 時点では教育漢字の大半が 90%に達し、ほぼ定着したと見ることができる。

3. 2 品詞別の漢字・仮名の書き分け

品詞別の漢字使用傾向を調査する。最初に、品詞別 (自立語のみ) の 1 万語あたりの語彙頻度を示す。

【表 5】学年別の品詞使用頻度 (1 万語あたり・自立語のみ)

| 品詞 | 小 1 | 小 2 | 小 3 | 小 4 | 小 5 | 小 6 | 中 1 | 中 2 | 中 3 |
|-----|------|------|------|------|------|------|------|------|------|
| 名詞 | 1882 | 1801 | 2043 | 1954 | 2015 | 2089 | 2112 | 2022 | 2182 |
| 動詞 | 1229 | 1299 | 1291 | 1376 | 1380 | 1423 | 1431 | 1461 | 1449 |
| 代名詞 | 198 | 247 | 224 | 221 | 214 | 229 | 235 | 266 | 236 |
| 副詞 | 283 | 276 | 236 | 236 | 244 | 214 | 206 | 217 | 187 |
| 形容詞 | 210 | 173 | 173 | 173 | 172 | 168 | 171 | 180 | 178 |
| 形状詞 | 125 | 118 | 111 | 108 | 121 | 126 | 132 | 135 | 132 |
| 連体詞 | 46 | 55 | 65 | 78 | 91 | 101 | 105 | 111 | 108 |
| 接続詞 | 28 | 38 | 33 | 46 | 47 | 51 | 53 | 51 | 53 |
| 感動詞 | 109 | 47 | 25 | 32 | 20 | 14 | 11 | 10 | 6 |

次に、これらの品詞について、漢字使用率の学年別推移を以下に示す³。

² 学習指導要領の中学 3 年次の文字に関する事項に「学年別漢字配当表に示されている漢字について、文や文章の中で使い慣れること。」とあるので、一応の目安とした。

³ 漢字使用率は、品詞別の漢字書き数/当該品詞数 (漢字書き+仮名書き) で集計した。出現形が一字でも漢字を含んでいる場合、漢字書きと判定した。例えば「名まえ」のような表記も、漢字書きと判定した。以下の調査も同様に処理した。

【表6】学年別の品詞別漢字使用率 [単位: %]

| 品詞 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 |
|-----|-----|------|------|------|------|------|------|------|------|
| 名詞 | 9.5 | 32.9 | 50.9 | 56.5 | 67.1 | 74.0 | 77.8 | 77.8 | 78.6 |
| 動詞 | 1.5 | 12.5 | 19.6 | 25.7 | 29.0 | 34.4 | 41.2 | 42.1 | 42.7 |
| 代名詞 | 1.6 | 1.5 | 18.3 | 26.8 | 37.4 | 49.6 | 60.8 | 61.5 | 64.1 |
| 形容詞 | 2.4 | 4.4 | 9.2 | 10.3 | 12.9 | 19.3 | 22.1 | 23.7 | 24.0 |
| 形状詞 | 3.7 | 14.4 | 25.0 | 30.8 | 33.1 | 45.3 | 52.9 | 53.5 | 56.5 |
| 副詞 | 2.6 | 13.0 | 19.1 | 28.6 | 38.0 | 43.3 | 47.8 | 52.1 | 52.0 |
| 連体詞 | 2.5 | 6.1 | 9.0 | 8.4 | 7.1 | 8.1 | 10.7 | 12.5 | 9.2 |
| 接続詞 | 0.0 | 0.6 | 0.0 | 0.0 | 3.1 | 3.5 | 2.2 | 2.2 | 2.6 |
| 感動詞 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4 | 0.0 | 0.0 | 0.0 |

品詞別に漢字書きの比率は異なり、およそ体言>用言>その他の語の順で漢字書きの比率が高くなる傾向がある。中3の比率を分母として百分率を計算したものを以下に示す。

【表7】学年別の品詞別漢字使用率（中3を100%として） [単位: %]

| 品詞 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 |
|-----|------|------|------|-------------|-------------|-------------|------|-------|-------|
| 名詞 | 12.0 | 41.9 | 64.8 | <u>71.9</u> | 85.4 | 94.1 | 99.0 | 99.0 | 100.0 |
| 動詞 | 3.5 | 29.3 | 45.9 | 60.3 | 68.0 | <u>80.6</u> | 96.5 | 98.7 | 100.0 |
| 代名詞 | 2.4 | 2.4 | 28.6 | 41.8 | 58.3 | <u>77.3</u> | 94.8 | 95.8 | 100.0 |
| 形容詞 | 10.2 | 18.2 | 38.2 | 43.0 | 53.7 | <u>80.2</u> | 92.0 | 98.8 | 100.0 |
| 形状詞 | 6.5 | 25.6 | 44.3 | 54.4 | 58.6 | <u>80.2</u> | 93.7 | 94.7 | 100.0 |
| 副詞 | 5.0 | 25.0 | 36.7 | 55.1 | <u>73.2</u> | 83.3 | 91.9 | 100.2 | 100.0 |

漢字書きの浸透が最も早いのは名詞で、小4年次で70%を超える。動詞、代名詞、形容詞、形状詞、副詞は小学校高学年の段階で70%に到達する。中学1年次には、全ての品詞の漢字使用率が90%を超えるが、この理由として中学生になれば漢字で書ける語は品詞に関わりなく漢字で書くという意識の変化(または、教師の指導)があると考えられる。

3. 3 語種別の漢字・仮名の書き分け

語種別の漢字使用傾向を調査する。最初に、語種別の1万語あたりの語彙頻度を示す。

【表8】学年別の語種使用頻度（1万語あたり・記号など除く）

| 語種 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 |
|------|------|------|------|------|------|------|------|------|------|
| 和語 | 7172 | 6946 | 6906 | 7021 | 6906 | 6981 | 7047 | 7158 | 7095 |
| 漢語 | 1010 | 1044 | 1172 | 1106 | 1215 | 1282 | 1369 | 1269 | 1463 |
| 外来語 | 170 | 222 | 259 | 237 | 228 | 195 | 170 | 151 | 156 |
| 混種語 | 147 | 115 | 102 | 83 | 97 | 97 | 102 | 105 | 101 |
| 固有名詞 | 47 | 47 | 60 | 46 | 57 | 49 | 35 | 29 | 29 |

次に、これらの語種について、漢字使用率の学年別推移を以下に示す。

【表 9】 学年別の語種別漢字使用率 (記号など視く) [単位: %]

| 語種 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 |
|------|-------------|-------------|------|------|------|------|------|------|-------------|
| 和語 | 1.4 | 6.8 | 11.9 | 13.9 | 15.6 | 18.2 | 19.8 | 20.2 | 19.8 |
| 漢語 | 12.3 | 36.0 | 57.1 | 66.0 | 79.2 | 86.5 | 90.7 | 92.0 | 93.1 |
| 外来語 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 |
| 混種語 | 1.6 | 24.7 | 21.7 | 33.0 | 45.9 | 62.3 | 80.9 | 88.1 | 87.9 |
| 固有名詞 | 9.5 | 22.6 | 28.7 | 42.3 | 49.0 | 55.6 | 62.6 | 55.1 | 60.5 |

漢語は漢字との関連性が高い。本調査でも、中3の段階では93.1%の漢語が漢字書きされている。そこで、子ども達が作文時にある語(漢語)を思いついても漢語は漢字で書くものだという規範意識が働いて、当該の漢語の使用をひかえるのではないかという予想ができる。しかし実際には、学習した漢字が少ない低学年においても、仮名書きの漢語が多数使用されている(小1:12.3%、小2:36.0%)。予想とは異なり、漢字で書くべきという規範意識の語彙の選択への影響は低いと考えられる。中3の比率を分母として百分率を計算したものを以下に示す。

【表 10】 学年別の語種別漢字使用率 (中3を100%として) [単位: %]

| 語種 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 |
|------|------|------|------|-------------|------|-------------|-------|-------|-------|
| 和語 | 7.2 | 34.4 | 60.3 | 69.9 | 78.9 | 91.9 | 100.1 | 102.0 | 100.0 |
| 漢語 | 13.2 | 38.7 | 61.4 | 71.0 | 85.1 | 93.0 | 97.4 | 98.8 | 100.0 |
| 混種語 | 1.9 | 28.1 | 24.7 | 37.5 | 52.2 | 70.9 | 92.0 | 100.3 | 100.0 |
| 固有名詞 | 15.7 | 37.3 | 47.4 | 69.9 | 80.9 | 91.9 | 103.5 | 91.1 | 100.0 |

漢語、和語、固有名詞は小4年次でほぼ70%に達し、混種語は小6年次で70%に達する。中1の段階では、いずれも90%を超える。

3. 4 高頻度語彙における漢字使用頻度

使用頻度の高い動詞20種について漢字の使用実態を調べる。最初に、それらの動詞の100万語あたりの使用頻度を示す。順位は全学年の平均頻度による。

【表 11】 学年別の語彙使用頻度(動詞頻度上位20語/100万語あたり)

| 語彙素 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 | 平均 |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 為る | 13485 | 18657 | 15482 | 16706 | 17143 | 17538 | 17744 | 17429 | 18625 | 16979 |
| 居る | 6940 | 9943 | 12435 | 12794 | 12960 | 14478 | 13181 | 14710 | 12913 | 12262 |
| 成る | 14963 | 14053 | 11512 | 11101 | 10908 | 10160 | 9975 | 8780 | 9770 | 11247 |
| 言う | 3817 | 7078 | 7166 | 7007 | 6980 | 7814 | 7090 | 8790 | 7799 | 7060 |
| 思う | 2897 | 4967 | 5754 | 7472 | 7715 | 8345 | 9243 | 8838 | 7809 | 7005 |
| 有る | 4527 | 4021 | 4893 | 5064 | 5945 | 7192 | 6636 | 7497 | 8002 | 5975 |
| 頑張る | 9834 | 6657 | 5942 | 4305 | 4904 | 4560 | 4601 | 4544 | 4137 | 5498 |
| 行く | 3318 | 3658 | 3845 | 4048 | 4681 | 4513 | 4436 | 4421 | 4746 | 4185 |
| 出来る | 3586 | 4013 | 3047 | 3553 | 3481 | 3661 | 4301 | 3861 | 4498 | 3778 |
| 見る | 4384 | 2689 | 3110 | 4918 | 3264 | 2807 | 3282 | 3169 | 3435 | 3451 |
| 遣る | 3247 | 3938 | 3342 | 3277 | 2899 | 2808 | 2123 | 2756 | 1482 | 2875 |
| 来る | 2486 | 2476 | 2016 | 2622 | 2548 | 1926 | 2410 | 2665 | 2402 | 2395 |
| 仕舞う | 519 | 1007 | 1461 | 2011 | 2266 | 1924 | 1746 | 2176 | 1655 | 1641 |

| | | | | | | | | | | |
|-----|------|------|------|------|------|------|------|------|------|------|
| 作る | 3573 | 2404 | 2077 | 1478 | 1119 | 1242 | 1052 | 677 | 810 | 1604 |
| 考える | 292 | 287 | 667 | 634 | 954 | 1667 | 1442 | 1817 | 2107 | 1096 |
| 呉れる | 540 | 1037 | 1038 | 1197 | 1536 | 1178 | 1303 | 1050 | 939 | 1091 |
| 貰う | 1541 | 1401 | 1516 | 1167 | 925 | 822 | 815 | 691 | 500 | 1042 |
| 出る | 916 | 1174 | 1074 | 1340 | 1101 | 940 | 934 | 988 | 902 | 1041 |
| 分かる | 434 | 564 | 821 | 839 | 1095 | 1259 | 1272 | 1211 | 1064 | 951 |
| 入る | 519 | 612 | 1091 | 950 | 875 | 1015 | 1216 | 1124 | 968 | 930 |

次に、これらの動詞について、漢字使用率の学年別推移を以下に示す。表の再右列は辞書形に含まれる漢字の配当学年(複数の漢字が含まれる場合は最も低い学年)である。

【表12】 学年別の語彙別漢字使用率(動詞頻度上位 20 語) [単位: %]

| 語彙素 | 小1 | 小2 | 小3 | 小4 | 小5 | 小6 | 中1 | 中2 | 中3 | 配当 |
|-----|------|------|------|------|------|------|------|-------|------------|-----|
| 為る | 0.0 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | <u>0.0</u> | 常用 |
| 居る | 0.6 | 0.0 | 0.0 | 0.0 | 0.1 | 0.2 | 0.0 | 0.1 | <u>0.1</u> | 小5 |
| 成る | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 | 0.0 | 0.0 | <u>0.1</u> | 小4 |
| 言う | 1.0 | 27.5 | 28.9 | 26.2 | 19.4 | 22.9 | 24.7 | 27.6 | 21.0 | 小2 |
| 思う | 10.3 | 88.5 | 92.5 | 97.1 | 98.6 | 99.5 | 99.8 | 99.6 | 99.9 | 小2 |
| 有る | 3.3 | 0.0 | 0.5 | 0.0 | 1.0 | 0.1 | 0.0 | 0.0 | <u>0.2</u> | 小3 |
| 頑張る | 2.3 | 1.1 | 0.0 | 0.4 | 9.3 | 32.3 | 59.0 | 68.9 | 73.1 | 小5 |
| 行く | 0.0 | 19.5 | 31.8 | 36.8 | 28.9 | 33.7 | 24.3 | 26.1 | 29.3 | 小2 |
| 出来る | 0.8 | 2.4 | 11.0 | 11.6 | 7.1 | 13.9 | 13.9 | 12.6 | 10.3 | 小1 |
| 見る | 7.5 | 46.8 | 41.9 | 43.9 | 46.1 | 51.6 | 63.4 | 57.1 | 65.2 | 小1 |
| 遣る | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.3 | 0.0 | <u>0.0</u> | 常用 |
| 来る | 0.9 | 3.8 | 11.3 | 15.1 | 15.0 | 12.4 | 17.2 | 14.6 | 11.0 | 小2 |
| 仕舞う | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | <u>0.0</u> | 小3 |
| 作る | 9.4 | 46.5 | 65.5 | 61.3 | 78.6 | 82.9 | 54.9 | 66.2 | 68.7 | 小2 |
| 考える | 0.0 | 79.5 | 91.7 | 96.0 | 99.2 | 99.7 | 99.7 | 100.0 | 100.0 | 小2 |
| 呉れる | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | <u>0.0</u> | 常用 |
| 貰う | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.9 | 2.6 | <u>0.0</u> | 常用外 |
| 出る | 6.8 | 55.0 | 49.7 | 61.4 | 62.1 | 69.4 | 78.4 | 81.4 | 83.4 | 小1 |
| 分かる | 0.0 | 15.9 | 25.9 | 40.5 | 54.4 | 54.8 | 67.2 | 70.5 | 71.9 | 小2 |
| 入る | 0.0 | 74.0 | 83.4 | 87.8 | 90.6 | 98.1 | 99.5 | 98.9 | 99.0 | 小1 |

これらの動詞の多くは中学校までに習う漢字(常用漢字)で漢字書きが可能だが、中3段階でもほとんど漢字書きされないもの、漢字と仮名の書き分けがあるもの、ほとんど漢字書きされるものがあることが分かる。ほとんど漢字書きされない語(中3年次で漢字書きが1%未満の語に下線を付した)以外の語の多くは、使用される漢字が小1～小2に配当されており、高頻度語でありながら学習時期が遅い漢字は表の範囲では見当たらない。その意味で教育漢字の配当順は、子ども達の使用実態に即したものであると評価することができる。「頑張る」が唯一の例外だが、これは「がんばったこと」という課題の影響で頻度が高くなっているだけであり、本来はそれほど高頻度の語ではないと考えられる。

4. 大人の文章との対照

4. 1 横断的分析: 作文の文体的特徴

作文の文体的特徴を確認するために、大学生の作文⁴(「夢」「がんばったこと」と BCCWJ コアデータ(知恵袋、ブログ、書籍、雑誌、新聞、白書)の漢字使用頻度を、漢字種別(小学校配当の教育漢字、教育漢字以外の常用漢字、常用外漢字)ごとに調べた。結果を以下に示す。

【表13】レジスタ別の漢字使用頻度(1万字あたり)

| レジスタ | 配1年 | 配2年 | 配3年 | 配4年 | 配5年 | 配6年 | 常用 | 常用外 | 計 |
|-------|-------|-------|--------|-------|-------|-------|-------|------|--------|
| 白書 | 528.4 | 884.0 | 1038.6 | 732.8 | 748.4 | 296.1 | 401.8 | 3.7 | 4633.8 |
| 新聞 | 779.6 | 912.6 | 745.7 | 544.0 | 453.5 | 231.1 | 414.0 | 32.1 | 4112.7 |
| 雑誌 | 542.9 | 641.6 | 528.2 | 348.3 | 274.9 | 155.7 | 354.2 | 47.0 | 2892.6 |
| 書籍 | 547.5 | 653.8 | 526.0 | 325.1 | 257.1 | 159.1 | 318.9 | 37.4 | 2825.1 |
| 作文(夢) | 567.5 | 684.8 | 462.3 | 309.7 | 368.3 | 166.0 | 216.1 | 21.3 | 2796.0 |
| 作文(が) | 587.8 | 715.3 | 499.6 | 319.6 | 213.5 | 124.8 | 198.4 | 12.3 | 2671.3 |
| ブログ | 458.2 | 633.6 | 453.5 | 303.7 | 201.0 | 125.6 | 285.8 | 36.7 | 2498.2 |
| 知恵袋 | 413.7 | 641.8 | 427.8 | 299.9 | 209.6 | 132.8 | 265.5 | 19.6 | 2410.6 |

大学生の作文の漢字使用頻度は1万字あたり2671~2796字である。4000字以上である白書・新聞とは大きな隔りがあるが、それ以外のレジスタとは極端な差はなく、おおよそ雑誌・書籍とブログ・知恵袋の中間程度である。

「夢」と「がんばったこと」では「夢」の漢字使用頻度が高く、特に5年配当漢字の使用頻度が高い。これは5年配当である「夢」という漢字が多く含まれている(1万字あたり約160字)ためであり、それを除外すると「夢」と「がんばったこと」の差は小さくなる。それ以外の特徴としては、「夢」の方が6年配当や配当外の漢字使用頻度が高く、「がんばったこと」の方が低~中学年配当の漢字使用頻度が高い傾向がある。これは2つの課題で使用される語彙の違いを反映している可能性があるが、より詳細な分析は今後の課題としたい。

4. 2 縦断的分析: 漢字使用能力の飽和状況

中学生までの漢字学習で、児童・生徒の漢字使用能力がどの程度まで大人の漢字使用能力に接近するかを見るために、中学3年次の漢字使用実態と大学生の漢字使用実態を対照する。配当学年別の漢字使用頻度、品詞別、語種別の漢字使用率について、中3と大学生を対照した表を以下に示す。それぞれ表3、6、9の中3の数値に、大学生の数値を並べたものである。

【表14】配当学年別の漢字使用頻度(1万字あたり)

| 配当漢字 | 配1年 | 配2年 | 配3年 | 配4年 | 配5年 | 配6年 | 常用 | 常用外 | 計 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 中3 | 646 | 705 | 518 | 320 | 180 | 127 | 223 | 6 | 2726 |
| 大学生 | 588 | 715 | 500 | 320 | 213 | 125 | 198 | 12 | 2671 |

⁴ 大学生1, 2年生に調査協力を依頼し、「ゆめ」「頑張ったこと」でそれぞれ作文課題を課した。これにより、「ゆめ」108編、「頑張ったこと」223編の作文を収集した。なお調査に際して、A4用紙1枚程度(1600字)という目安を示したが自宅での課題としたため、条件に幅があることを断っておく。

【表 1 5】品詞別の漢字使用率 [単位: %]

| 品詞 | 名詞 | 動詞 | 代名詞 | 形容詞 | 形状詞 | 副詞 | 連体詞 | 接続詞 | 感動詞 |
|-----|------|------|------|------|------|------|------|-----|-----|
| 中3 | 78.6 | 42.7 | 64.1 | 24.0 | 56.5 | 52.0 | 9.2 | 2.6 | 0.0 |
| 大学生 | 67.8 | 42.0 | 60.4 | 25.6 | 58.2 | 52.9 | 11.9 | 1.3 | 0.0 |

【表 1 6】語種別の漢字使用率 [単位: %]

| 語種 | 和語 | 漢語 | 外来語 | 混種語 | 固有名詞 |
|-----|------|------|-----|------|------|
| 中3 | 19.8 | 93.1 | 0.0 | 87.9 | 60.5 |
| 大学生 | 19.8 | 78.6 | 0.7 | 85.8 | 88.9 |

個別の項目を見ると、中3の時点ではほぼ大学生と同等の水準に達しているもの、大学生の水準にやや及ばないもの、中3の時点の方がむしろ数値が高いものがある。例えば表14は、個々の項目について前後はあるが、全体としては中3の方が大学生より漢字使用頻度が高いことを示している。表15を見ると、名詞、動詞、代名詞、接続詞は中3の方が漢字使用率が高いが、形容詞、形状詞、連体詞は大学生の方が漢字使用率が高い。表16を見ると、漢語や混種語は中3の方が漢字使用率が高いが、固有名詞は大学生の方が漢字使用率が高い。これらの差異の意味を分析するためには、各項目に含まれるどのような語彙が差異を生み出しているかについて、より詳しく調査する必要がある。しかし全体としては、これらの数値は概ね中学3年次の漢字使用能力が大学生の漢字使用能力に接近していることを示しており、高校以降の変化が無いとまでは言えないものの、中学修了段階でかなりの程度飽和状態に近づいていると考えられる。

5. おわりに

本発表では、作文コーパスに基づいて児童・生徒の漢字使用能力の推定を試みた。また、BCCWJのコアデータや大学生の作文と対照することによって、子ども達の漢字使用能力が大人のそれにどの程度近づいているのかについても言及し、発達過程の概要を示した。

より詳細な分析を進めるために、現在『児童・生徒作文コーパス』の内10万形態素程度を目標に(全体の1割弱)、自動解析後に人手修正を行ったコアデータの構築を進めている。現在使用しているデータは自動解析によって形態論情報等を付与しているが、誤字脱字や仮名書きが多い低学年の作文は自動解析の精度が低く、十分な信頼性を確保できていない。人手で形態論情報、構文情報を付与したコーパスを整備することによって、本発表で得られた調査結果を再検討するとともに、今後は以下のような課題の分析を進めていきたい。

- ・ 同一語での仮名書きと漢字書きの傾向差に関する議論
- ・ 同一漢字を用いる異語の漢字書きの傾向差に関する議論（「下る」と「下がる」など）
- ・ 作文文型の発達と語彙・漢字使用についての議論

本研究は『児童・生徒作文コーパス』を使用した一連の研究の一部である。これと並行して、発表者ら以外の共同研究者によって同コーパスを利用した作文能力の発達過程の推定と数値化が進められている。中でも子ども達の漢字使用能力に関する研究は、現場からの要請が強く、率先して進められるべきものの一つである。本研究の最終的な目標は、教育現場における作文教育の改善と適正化を図ることにある。研究が進み、言語研究の立場から現場の教師が手軽に利用できる漢字使用の実態の分析や作文指導の指針を提案し、有

効に活用されれば、昨今二者の乖離が叫ばれて久しい研究と教育の現場の協働の一つの形として位置づけることができる。

謝 辞

本研究は、平成 27 年度 漢字・日本語教育研究助成制度「作文コーパスを資料に児童・生徒の漢字使用・選択傾向と発達の実態を明らかにする。一語彙情報つき作文コーパスの構築と学齢別語彙・漢字使用実態調査」(研究代表者：宮城信)、および日本学術振興会科学研究費補助金基盤研究(B)「作文を支援する語彙・文法的事項に関する研究」(平成 26～30 年度、研究代表者：矢澤真人、研究課題番号：26285196)による補助を得ています。

文 献

河内昭浩(2015)「国語教育のための「常用漢字表」語例の検討」『第 7 回コーパス日本語学ワークショップ予稿集』 pp.113-122.(https://www.ninjal.ac.jp/event/specialists/project-meeting/files/JCLWorkshop_no7_papers/JCLWorkshop_No7_web.pdf よりダウンロード可能)

鈴木一史(2011)「作文コーパスからみる生徒の使用語彙」『特定領域「日本語コーパス」平成 22 年度公開ワークショップ(研究成果報告会)予稿集』 pp.343-350.(http://www.ninjal.ac.jp/corpus_center/bccwj/doc/workshop/JC-G-10-02.pdf よりダウンロード可能)

田中牧郎(2011)「語彙レベルに基づく重要語彙リストの作成 ―国語政策・国語教育での活用のために―」 pp.77-87.(http://www.ninjal.ac.jp/corpus_center/bccwj/doc/workshop/JC-G-10-02.pdf よりダウンロード可能)

宮城信・今田水穂(2015)「『児童・生徒作文コーパス』の設計」『第 7 回コーパス日本語学ワークショップ予稿集』 pp.223-232.(https://www.ninjal.ac.jp/event/specialists/project-meeting/files/JCLWorkshop_no7_papers/JCLWorkshop_No7_web.pdf よりダウンロード可能)

関連 URL

作文を支援する語彙文法的事項に関する研究プロジェクト <https://sites.google.com/site/sakubunshienproject/>