

受容する文章と産出する文章の難易度比較

安部朋世（千葉大学）・橋本修（筑波大学）

1 リーダビリティ研究と（日本の）国語教育の流れ

母語を含めた言語学習において、「このレベルの学習者にこの難度の教材は適切なのか」等、文章の難易度が問題になることは多い。このことは言語の学習・教育においては歴史上常に問題になるはずであるが、教育目的の難易度研究が本格的におこなわれたのは 1950 年以降であると見られる。

まず、リーダビリティ周囲の研究として、国立国語研究所 1951a「国語学力標準設定に関する調査研究」の中で（日本語母語話者における）読解力・語彙力の標準的測定の必要性が述べられ、国立国語研究所 1951b「義務教育終了者に対する語彙調査の試み」ではそれにもとづいた、インフォーマントにおける語彙獲得の状況が実際に調査されている。調査の概略は高校 1 年生 15 名において、37970 語についての獲得状況を質問紙法で回答してもらうというもので、事実上語彙の難易度計測のひとつの手法として捉えられるものである。

また、阪本一郎 1952「文章の語彙比重の査定法—Readability の研究の一つの試み—」は 3 段階の難度に分けた語彙表をもとに各種文章の難度を計測し、国語教育において Readability を数値で表現する嚆矢の研究である。

そののちしばらくの期間を置いて、より工学的な手法・問題意識から Readability 測定の各種手法が発達したのは、概ね 2000 年代に入ってからである（注 1）。日本語教育の隆盛による養成と、計測の手段・間接材料となるコーパスの発達が主な背景にあると思われる。現在、利用のしやすい Readability 計測ツールとして、以下の 3 つが挙げられる。

「日本語読解学習支援システム リーディングチュウ太」

「obi（帯）3」

「日本語文章難易度判別システム（jReadability）」

それぞれの詳細については、川村よし子・北村達也ほか 2008、佐藤理史 2011、柴崎秀子・原信一郎 2010、李在鎬 2016 などを参照。3 種のリーダビリティ測定器にはそれぞれある程度異なった特徴があり、必要に応じて使い分ける必要がある。

2 各種 Readability 測定値の利用について

2-1

リーダビリティ測定器により産出された値の利用のうち、最も普通なのは

- 1) 学習者が受容する(学習者に受容させる)文章の難易度を測り、教材選択の参考にする」
- 2) 学習者が受容する文章中の語彙の難易度を計っておき、指導の参考にする」というものである。

2)については、「分からない語については学習者が辞書を引く」ということで解決すると

いうケースもあるが、難度の高い、分からない語の数が一定程度を越えると学習者がストレスに耐えきれず、一種の放棄をおこなう心配がある。逆に、受容する文章中のほとんどの語が既知で、1つか2つの語だけが難度の高い語彙で、その語についてのみ辞書を引けば印象が強くなり、定着が促進されることがあるかもしれない。そのような学習状況・受容状況を予測し準備する助けとしても有用である。

また、2)語彙の難易度情報の利用は、いわゆる教材には止まらない。教材にはないが、教員が使用する語彙の中に当該学習レベルに合わない語彙が紛れ込むリスクがあり、そのような語をピックアップして留意する、というような利用法も可能である。例えば教材そのものには出現しないが、「物語の展開に注意して読もう」という場合の「展開」という語をいつごろから使ってよいか、ということを検討できる（ただしこの点については国語教育においては特に「原理的にはできる」ということであって、現時点ではそこまで特定の検討は難しいかもしれない。ちなみに「展開」の難易度は「リーディングチュウ太」においてはN2N3、jReadabilityにおいては中級後半という判定である。後述するが、日本語教育を主目的としたこれらの難易度測定器を国語教育とどうすりあわせていくのかは、ひとつの今後の課題である）。

2-2

本発表者らはこれまで、国語教育においても readability という概念が国語教育にも有用であり、既存の難易度測定器を用いてできることについて検討してきた。ひとつは、児童・生徒の産出する作文群について jReadability によって難易度測定をおこない、直感的な児童・生徒作文の発達を確認することができた。橋本修・安部朋世・関口雄基 2018 の計測によれば、宮城信・今田水穂 2015 等で示されている小学生・中学生作文の、学年ごとの計測値は、

小学1年～3年生 jReadability 値 4.17 （分散 1.09）

小学4年～6年生 jReadability 値 3.86 （分散 0.42

中学生 jReadability 値 3.42 （分散 0.37）

となっている（jReadability 値は最小値約 0.4、最大値約 6.5 で、値が小さいほど難度が高い）。これは、当該児童・生徒の作文の水準がある程度順調に上がっていること、学年が上がるにつれて、産出する作文の質が（難易度という点では）均質化することを示している。また参考値として大学生の意見文、新聞社説についても同様の調査をおこない、

大学生意見文 jReadability 値 2.30 （分散 0.23）

新聞社説 jReadability 値 2.04 （分散 0.18）

という値を得ている。

語彙の難易度という側面については、「目標」「ことがら」等の論理語彙について、jReadability に付されている「初級前半（最も易しい）～級外（最も難しい）」という判定を 0～6 の間隔尺度と見なして数値化した。いっぽうで小学校 3 年生 33 名・6 年生 34 名に対して理解度アンケートをおこなって数値化し、両者の相関を見た。その結果は、「目標」「ことがら」「結果」「機能」「すじみち」「つくり」「構造」「原因」「くみたて」「理解」「まこと」「めあて」「めじるし」「真実」「ねらい」の 16 語について、jReadability 値と小学校 3 年生の理解度値の相関係数 0.62、jReadability 値と小学校 6 年生の理解度値の相関係数 0.62 となり、一定の相関があることが確認された。

3 受容する文章と産出する文章との難易度比較

3-1

これまでの検証とは別に、他にもリーダビリティ測定器の値を利用できる領域として、本発表では「受容する文章と産出する文章の難易度比較」をおこなう。国語学習者は受容する文章の一部を受け取って、自らの作文の水準を高めていくと考えられる。素朴に考えるなら、例えば小学校 1 年生は小学校 1 年生用教科書の文章を受容して、その中から身につけられた一部を使って作文をするというモデルを考えることもできるが、特に母語話者においては、学校教科書以外の文章を受容し、その影響を受けた文章作成能力を獲得しているケースも考えられる。また、学習者は受容した文章の全てをみずからのものにするわけではないであろうから、受容する文章と産出する文章には一定の差がある可能性が示唆される。このような考察を背景として、以下のような調査を開始した。大まかな手順は以下の通りである。

- a) 宮城信・今田水穂 2015 等で示された小学生・中学生作文を「obi3」によって難易度測定する。難易度スケールは obi3 搭載の T13（小学校 1 年生を難易度 1、中学校 3 年生を難易度 9、高校 3 年生を難易度 12、大学生・一般を 13 としたスケール）を採用する。
- b) それぞれの学年の作文の難易度を計測し、それぞれの学年難易度スケール値を比較する。

上記の方法をとったのは、obi3 の T13 スケールにおける難易度値が主として各学年の教科書を元にして算出されているためである。すなわち、極めて大まかに言えば、ある文章を obi3 の T13 スケールによって難易度測定するとき、その値は「〇〇学校〇年生の教科書と同じぐらいの難易度である」ということを示している。したがって、例えば小学校 4 年生の文章を上記の方法で難易度測定すると、「この作文は小学校〇年生の教科書と同じぐらいの難易度である」という結果が得られることになる。この方法により、現時点で小学校 1 年生～中学校 3 年生の作文各 15 について計測をおこなっている。

3-2

現時点で得られた値は概略下記のようなものである。

小学校 1 年生	測定値 2.07	T13 の学年教科書難易度 1	差分+1.07
小学校 2 年生	測定値 3.27	T13 の学年教科書難易度 2	差分+1.27
小学校 3 年生	測定値 4.60	T13 の学年教科書難易度 3	差分+1.60
小学校 4 年生	測定値 5.00	T13 の学年教科書難易度 4	差分+1.00
小学校 5 年生	測定値 6.13	T13 の学年教科書難易度 5	差分+1.33
小学校 6 年生	測定値 6.73	T13 の学年教科書難易度 6	差分+0.73
中学校 1 年生	測定値 7.47	T13 の学年教科書難易度 7	差分+0.47
中学校 2 年生	測定値 6.40	T13 の学年教科書難易度 8	差分-1.60
中学校 3 年生	測定値 7.67	T13 の学年教科書難易度 9	差分-1.33

測定値は各学年 15 名の平均値なので小数になるが、obi3 の T13 の難易度は「〇年教科書相当」となるので整数値である。サンプル数が充分ではないので憶測になる（注 2）が、「概

ね小学生においては、同一学年において児童作文のほうが教科書よりも難易度の高い文章を算出している」「概ね学年が上がるにつれて教科書の難易度のほうが難度上昇のピッチが高く、どこかのタイミング（ここでは中学校 2 年生）で（同一学年の）教科書文章のほうが難度が高くなる」という可能性が示唆される。当日資料ではデータ数を増やし、より詳細な検討をおこなう。

4 今後の課題

今後の課題としては以下のようなものが挙げられる。

- ・3・2 で示唆された傾向は、作文資料のジャンルが「夢」であることに大きく影響を受けている。すなわち、学年が上がるにつれて教科書文章のほうが難度が高くなるのは、教科書文章に論説文等が学年の上昇につれ頻出することによっている可能性がある。この点を明らかにするには、小学校教科書の文章をジャンル分けして測定・比較をおこなう必要がある。

- ・すでに若干触れたが、母語においては受容する文章が、教科書以外に豊富に与えられている場合が多く、その影響を検討する必要がある。外国語学習、非母語としての日本語学習の場合にどのような様相を見せるのか調査を今後おこないたい。

注

1 一時期リーダビリティ研究が少なかったのはリーダビリティという概念が本質的規定を受けつけにくいということがあるとも言われる。本発表者は「疑似的計測でも、計測するほうが計測しないよりよい」「計測の（それなりの）妥当性は、本発表や関連先行研究による計測結果とその考察から少なくとも間接的に妥当性が示されている」という立場をとる。

2 小学校 1 年生の作文の数値については特に問題がある。各作文において測定される難易度の最低値は 1 で、それ以上易しい（「0」等）1 未満の値がないため、本来の難易度よりも、数値が高くなる（難しめという数値が出る）という可能性がある。現時点ではこの点についてのよい解決策は見つかっていない。

参考文献

李在鎬 2016 「日本語教育のための文章難易度研究」『早稲田日本語教育学』 21

川村よし子・北村達也ほか 2008 「単語親密度と頻度情報を活用した難易度判別システム」

ppt 資料 <http://language.tiu.ac.jp/about.htm#papers>

佐藤理史 2011 「均衡コーパスを規範とするテキスト難易度測定」『情報処理学会論文誌』 52-4

柴崎秀子・原信一郎 2010 「12 学年を難易尺度とする日本語リーダビリティ判別式」『計量国語学』 27-6

橋本修・安部朋世・関口雄基 2018 「日本語リーダビリティ研究の、国語教育への応用可能性」『日中言語研究と日本語教育』 11

宮城信・今田水穂 2015 「『児童・生徒作文コーパス』の設計」『国立国語研究所第 7 回コーパス日本語学ワークショップ予稿集』