

2. 児童を対象とした単語親密度実験

A Survey on Word Familiarity of Elementary School Students

あらまし 児童を対象とした漢字書き取り調査の結果から、児童の語彙範疇外の説明語を使うことが、元の漢字の想起を妨げる大きな要因であることをこれまでに解明してきた。次の研究行程として、児童の語彙の範囲を調べるため、学習基本語彙表をもとに教科書初出学年ごとにまとめた単語の音声刺激を使って、言葉の親密度調査を行った。その結果、初出学年により単語親密度の平均は有意に異なった。高い親密度をもつ単語の割合は、初出学年が低い単語群ほど高く、学年が上がるにつれて低くなることを確認した。

1. はじめに

詳細読みの利用者が児童の場合の問題を解明するため、小学5年生の児童を対象とした漢字書き取り実験を既に実施した。その結果、児童の語彙範疇にないと思われる単語を説明表現に使用している問題の影響が大きいことがわかった[1]。これより、次の研究課題は、児童の語彙範疇をより厳密に特定することとなる。ここで重要なのは、児童は語彙の獲得期にあるので学年ごとの差が大きいと予測されることである。そこで、学習基本語彙表[2]を元に教科書初出学年で分類した単語群を刺激として、児童を対象とした言葉の親密度調査を行った。

2. 語彙の分類と選択

2. 1 小学生の語彙の分類

甲斐によると、小学生の語彙体系は、(1)基礎語彙、(2)学習基本語彙、(3)学習語彙、(4)一般語彙に整理される[2]。番号の大きい語彙はそれぞれ下位の語彙を含む(図1)。それぞれの内容と語数を以下に示す。

- (1) 基礎語彙：言語生活の上で最低限欠かせない1000～2000語。
- (2) 学習基本語彙：(1)を含む約5000語。小学生が表現活動に十分に駆使できる語彙。小学校の教科書は学習基本語彙を頻出させる。
- (3) 学習語彙：小学生用の国語辞典に登録されている語彙約25000語。語彙数は、『新教育基本語彙』[3]などを根拠とする。小学生の理解語彙の上限を示すとされる。
- (4) 一般語彙：小学生には直接には関係しない難解な語彙。中学生以上の国語辞典に登録されている語彙だが、小学校高学年の教科書に現れることもある。(3)と(4)の境界は明確ではない。

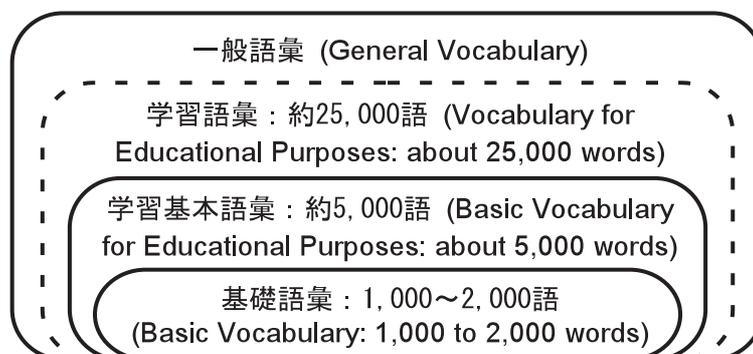


図1 小学生の語彙体系 (甲斐, 2002 を改変)

※参考：辞書の語彙数

(1) 子ども向け国語辞典とその見出し語数

- こどもこくごじてん（三省堂）： 1219 語（小学校 1・2・3 年生を対象）
- 学習国語新辞典第 6 版（小学館）： 約 19000 語
- 下村式 小学国語辞典（偕成社）： 24927 語
- チャレンジ小学国語辞典第 4 版（Benesse）：約 25000 語
- 例解小学国語辞典第 3 版（三省堂）： 約 33000 語

(2) 一般向け国語辞典とその見出し語数

- 現代国語辞典（学研）第 3 版： 約 67000 語
- 明鏡 国語辞典（大修館書店）： 約 70000 語
- 新明解国語辞典第 6 版（三省堂）：約 76500 語
- 国語辞典第 9 版（旺文社）： 約 81500 語
- 広辞苑第 5 版（岩波書店）： 約 230000 語

2. 2 語彙集の選択

児童の理解語彙の上限が学習語彙であり、その中でも児童が十分に駆使できるとされるのが学習基本語彙である。従って、説明表現としては学習基本語彙の中の単語を使用するのが適切と考えられる。なお、学習基本語彙より基本的な基礎語彙を説明語の使用限界としないのは、これだけでは教育漢字 1006 字を説明するには十分ではないという理由による。

比較的最近にまとめられ、入手可能な学習基本語彙として『語彙指導の方法 [語彙表編]』（2002）の中の「学習基本語彙一覧表」がある[2]。その語彙数は 4113 語で、語彙選定の観点として、(a)教科書の語彙分析、(b)読み物などの語彙分析、(c)理解語彙の調査、(d)使用語彙の調査、(e)学識者による選定、の 5 種類を取り入れたとしている。この語彙表では、光村図書の国語教科書における各単語の初出情報（学年、上下巻の別、ページ）が記されている。この情報は、学年ごとの語彙の違いを見るのに利用できる。今回の調査では、この語彙表をもとに刺激を作成した。

3. 親密度調査の実施

小学5年配当の教育漢字を含む単語を学習基本語彙から選び、その親密度が各単語の初出学年により異なるかどうかを調べる調査を小学5年生の児童を対象に行った。なお、この調査結果は、実際の詳細読み説明表現用単語選定のためのデータとしても利用する。

3. 1 調査用単語群の選定

調査用単語の選定手順を以下に記す。

- (1) 小学5年生配当漢字185字を含む単語を「学習基本語彙一覧表」から抽出した。抽出した語数は602語。
- (2) 音声提示のため、同音異義語を削除した。同音異義語のチェックには学習語彙である「新坂本教育基本語彙」(19271語)[3]を辞書として使用した。同音異義語削除後の語数は370語。
- (3) 小学5年生配当漢字を2文字以上含む熟語は(1)の手順で重複して抽出されているのでこれを削除した(例:「非常」)。重複語句削除後の語数は332語。
- (4) 人の目によるチェックを行った。チェックの観点は以下の通りである。
 - 1) 小学校で学習する音読み・訓読みと異なる読み方をする語は排除した。例:「素人(しろうと)」。
 - 2) 意味的に対になっている語がある場合は、片方を削除した。例:「無意識」と「意識」の場合、「無意識」を削除して「意識」を残した。
 - 3) 自動詞・他動詞の両方がある場合は、片方を削除した。例:「燃やす」と「燃える」の場合、「燃える」を削除して「燃やす」を残した。
 - 4) (2)で同音異義語ありとされなかったものの、著者ら(渡辺, 大杉, 山口)が同音異義語ありと判断した語を削除した。例:「賛成」に対しては「酸性」が想起されたので、「賛成」を削除した。

人によるチェック後の語数は298語。

- (5) 国語教科書の語彙初出学年に基づき、298語を下の4群に分類した。
 - 1) 既習群(問題A):134語(初出4年生以下:62語, 初出5年生:72語)
 - 2) 未習群(問題B):164語(初出6年生:48語, 教科書出現なし:116語)
- (6) 既習群と未習群それぞれにおいて、ランダムな提示順序とした。

3. 2 音声刺激の作成と提示

問題番号に続けて、調査用単語を男性アナウンサーに読み上げてもらったものを収録

し、音声刺激とした。回答時間は、ある単語の読み上げ後、次の問題番号が読み上げられるまでの時間で、これは2.5秒に設定した。ただし、回答欄が次の段に移るときは回答時間を4秒と長くした。調査の趣旨と回答手順の説明も、同じアナウンサーの声で収録し、問題とともにカセットテープとCD-Rに録音した。

調査は調査対象校の教室で行った。試験時にはカセットテープとCD-RのいずれかをCDラジオカセットレコーダ（ケンウッド CDXA3S）で再生した。調査の趣旨の音声を聞かせながら、教室の後方座席の児童にも十分聞こえるように音量を調整した。

3. 3 調査対象者

横浜国立大学教育人間科学部附属横浜小学校と同鎌倉小学校の2校で、5年生2クラスずつに参加してもらい、クラスごとに既習群と未習群いずれかの問題を割り当てた。2校とも光村図書出版の国語教科書を使用している。調査は3月上・中旬に実施したので、調査対象者は5年生の学習を修了した者と見なせる。

参加者数は2校合計で162人だが、横浜小学校の帰国子女6人の回答結果は利用しないため、解析に用いる回答者数は156人となった。既習群と未習群の回答者数は両方とも78人であった。どのクラスも男女比はおよそ1:1である。

3. 4 親密度の評定法

単語の親密度調査としてはNTTの『日本語の語彙特性』が学術研究で広く引用されている[4]。この文献では親密度を1（低）から7（高）の7段階で成人被験者に評定させている。一方、文部省が1960年から数年にわたって実施した「児童・生徒の語い力の調査」では、よく知っていることば、だいたいわかることば、ぼんやりわかることば、知らないことばの4段階で回答させている[5]。

両文献の評定法を参考に検討したところ、まず7段階の評定尺度法で回答させるのは児童には難しいと考えた。次に、「だいたいわかる」と「ぼんやりわかる」の区別がつきにくいと考えられたため、最終的には次の3段階で回答させることとした。

ア よく知っている

イ だいたいわかる

ウ 知らない

それぞれの名義選択の目安を、以下のように調査対象者に説明した。

「ア よく知っている」は、知っていると自信をもって言える言葉、自分でも使っている言葉。

「イ だいたいわかる」は、聞いたことはあるけど、自分ではあまり使わない言葉。アを選ぶほどよく知っているとは自信をもって言えない場合、こちらを選ぶこと。

「ウ 知らない」は、聞いたこともない言葉。

4. 調査結果と考察

4. 1 初出学年条件による比較

各単語を「ア よく知っている」と答えた者の割合を本稿では単語の親密度とする。初出学年条件ごとに両校のデータを足し合わせて4群のデータを作った。各群の単語親密度の分布を表示したのが図2(a)~(d)である。既習群((a)と(b))は未習群((c)と(d))より分布が右に偏っている、つまり親密度が高い語の割合が大きい。4群に分散分析を適用するとF値は8.60となった。これは $F(3,294)=2.63$ (上側5%)より大きいので、初出学年条件により単語親密度の平均は有意に変化していると言える。親密度80%以上の単語の割合を比べてみると、初出4年以下,5年,6年,教科書出現なしの順に、82.3%,72.2%,66.7%,53.4%と順番に下がっている。逆にいえば、初出学年が上がるにつれて親密度が低い単語の割合が増加していった。

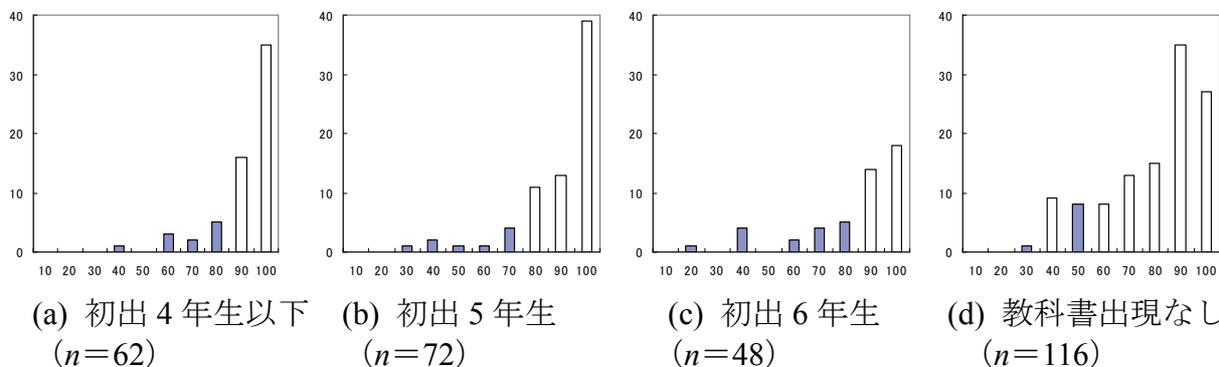


図2 親密度の度数分布。(a)~(d)いずれも横軸は単語親密度のデータ区間 [%], 縦軸は単語の語数。

4. 2 考察 - 詳細読み策定への示唆

学習基本語彙の単語であっても、初出学年が低い群ほど児童の単語親密度の高い単語の割合が大きかったことから、初出学年の低い単語を漢字説明表現に優先的に用いるべきだという示唆が得られた。

逆に、説明表現として選択を避けた方がよい単語の観点から考えると、例えば親密度50%未満の単語の個数と各群における割合は、初出4年以下では1個(割合:1.6%),5

年では4個（同5.6%）、6年では5個（同10.4%）と少ないが、教科書出現なしでは17個（同14.7%）まで増える。従って、教科書出現なしの群では、想起率が低い単語を選ぶ可能性が高まる。

5. まとめ

詳細読み説明表現用単語として、児童がよく知っている、あるいは知らない単語を特定するのが本研究の目的である。この目的のために利用可能な語彙集として「学習基本語彙」[2]を選んだ。次に、学習基本語彙に現れる単語でも初出学年により子どもの親密度が異なるのではと考え、5年生児童を対象に親密度調査を実施した。その結果、初出学年により単語親密度の平均は有意に異なった。高い親密度をもつ単語の割合は初出学年が低い単語群ほど高く、学年が上がるにつれて高い親密度をもつ単語の割合が低くなることを確認した。逆の観点から見ると、学年が上がるにつれ親密度の低い単語の割合が高くなることを示されている。これより、説明表現用単語の選択時には、単語の教科書初出学年を考慮する必要があると言える。特に、説明すべき漢字の配当学年より上の学年で初出する単語は、対象学年の子どもにとって親密度が低い語を含む割合が比較的高いので、単語の選定時には注意を要する。今回の調査で親密度が高った単語（例えば80%以上）の一覧は、詳細読み策定の際にそのまま用いることができる。同様に、親密度が低かった単語の一覧は、単語選定時に除外すべき単語として活用できる。

謝 辞

調査用カセットテープとCD-Rを御製作頂いた日本盲人会連合録音製作所の方々、調査に御協力頂いた横浜国立大学人間教育科学部附属横浜小学校及び鎌倉小学校の皆様
に感謝いたします。

参考文献

- [1] 渡辺哲也, 渡辺文治, 藤沼輝好, 大杉成喜, 澤田真弓, 鎌田一雄, “スクリーンリーダーの詳細読みの理解に影響する要因の検討—構成の分類と児童を対象とした漢字想起実験—,” 電子情報通信学会論文誌 D-I, Vol.J88-D-I, No.4, pp.891-899, April 2005.
- [2] 甲斐睦朗 (監), 語彙指導の方法 [語彙表編], 光村図書出版, 東京, 2002.
- [3] 国立国語研究所, 教育基本語彙の基本的研究, 国立国語研究所報告 117, 明治書院, 東京, June 2001.
- [4] 天野成昭, 近藤公久 (編著), NTT データベースシリーズ 日本語の語彙特性 第1期 CD-ROM 版 単語親密度, 三省堂, 東京, 2003.
- [5] 国立国語研究所, 日本語基本語彙—文献解題と研究—, 国立国語研究所報告 116, 明治書院, 東京, July 2000.

出典

本章は、以下の技術報告原稿をもとに再構成した。

- 渡辺哲也, 大杉成喜, 澤田真弓, 山口俊光, 渡辺文治, 岡田伸一: スクリーンリーダーの漢字詳細読みに関する研究—児童を対象とした言葉の親密度調査—, 電子情報通信学会技術報告, WIT2005-04, May 2005.