

教育基本語彙と成人の単語親密度との関係

山口 俊光 渡辺 哲也 大杉 成喜

国立特殊教育総合研究所

要旨

我々はスクリーンリーダの漢字詳細読みを視覚障害児に適用する際の問題について検討している。児童を対象とした詳細読みを作成する際には、児童の単語親密度を基に使用する単語を選択することが望ましい。しかしながら、児童を対象とした単語親密度の調査は、十分に行われているとは言い難い。そこで、本稿では過去に行った小規模な児童の音声単語親密度調査の結果を基に、成人の音声単語親密度から児童の音声単語親密度を予測可能かどうかを検討した。その結果、児童の音声単語親密度と成人の音声単語親密度の間には比較的強い相関が見られ、予測できる可能性が示唆された。

The Relationship of Word Familiarity between Children and Adults

Toshimitsu YAMAGUCHI Tetsuya WATANABE Nariki OSUGI

National Institute of Special Education

Abstract

We have been conducting a series of researches on the problem of *shosaiyomi* of kanji that arises when screen reader software is used by elementary school students with visual impairments. It is preferable that development of *shosaiyomi* that is used by elementary school students is based on word familiarity of children. However, there are few studies on word familiarity of children. In this paper, we have analyzed whether familiarity of elementary school students can be estimate from that of adults. As a result, a comparatively strong correlation was observed. Thus, it was suggested that familiarity of elementary school students can be estimate from familiarity of adults.

1 はじめに

1.1 背景

視覚障害者のコンピュータ利用場面では、画面に表示される文字、特に漢字を言葉だけで正しく伝えるための工夫がなされている。その工夫は、「詳細読み」と呼ばれ、漢字の音読みと訓読み、その漢字を含む熟語、漢字の構成要素(偏や旁)などを組み合わせた説明表現により、意図する漢字を一意にユー

ザに伝える。

この詳細読みの一部に、1度聞いただけでは元の漢字を想起しづらいものがあるという指摘がされたため [1] [2]、その分かりにくさの要因を調べ、これを改善して行くための研究を行っている。

研究は、詳細読みの利用者が大人の場合と子どもの場合とに分けて進めている。これは、大人と子どもの間で語彙の大きさに差があり、これが説明表現の理解に影響を及ぼすと推察されるためである。

この影響を明らかにするため、大人と子どもそれぞれについて、詳細読みを使用した漢字の想起実験を実施した。小学5年生を対象とした実験の結果、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の問題の影響が大きいことがわかった [3]。

さらに児童の語彙範疇をより正確に把握するため、児童を対象として単語 (298 語) の親密度調査を実施した [4]。この調査では、教科書における初出学年が低い単語群ほど親密度が高く、学年が上がるにつれて低くなることを確認した。

1.2 目的

本稿の目的は、成人を対象とした単語親密度 (約 8 万語) [5] と先の調査結果である 298 語の児童の単語親密度との関係を分析し、成人の単語親密度から児童の単語親密度の推測が可能であるかを検討することである。

児童にも分かりやすい詳細読みを作成するためには、児童の単語親密度に基づいて詳細読みの単語選択が行われることが理想的である。子どもを調査対象とした比較的規模の大きい語彙に関する調査として、文部省が 1957 年から 1962 年の間に実施した「児童・生徒の語い力調査」がある [6]。調査対象となったのは、小学校低学年児童、小学 4 年生、小学 6 年生、中学 1 年生、中学 3 年生である。調査対象ごとに調査語数にばらつきがあるが、約 2,000 語から 14,000 語について調査を行っている。しかし、調査対象の児童・生徒の学年が連続しておらず、網羅的な調査とは言い難い。また、調査から 40 年以上経過しており、現代の子どもの語彙力として使用するには適さないとされる。

児童の語彙について取り扱っている資料として、「学習基本語彙」(約 5,000 語) がある。この語彙に含まれる単語は、小学校教科書に頻出させることが望ましいとされている [7]。児童に分かりやすい表現にするため、我々は教育漢字の詳細読みで使用する単語を、できるだけこの学習基本語彙から選択している。しかし、教育漢字を含んだ語彙が網羅されているわけではないので、学習基本語彙から単語を選択できない場合がある。この際、学習基本語彙以外の語彙集から単語を選択する必要がある。

網羅的に成人の単語親密度を調査した資料として、『日本語の語彙特性』 [5] がある。調査語数は約 8 万語である。規模の面では、詳細読み作成の資料

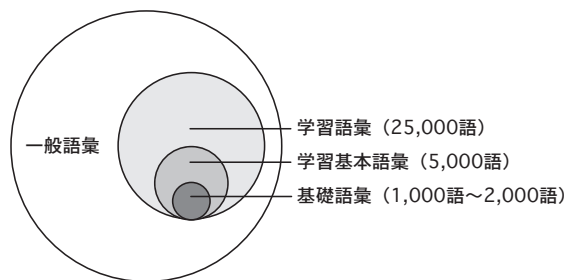


図 1: 語彙体系の相互関係

として十分なデータであるといえる。そこで、この成人の単語親密度から児童の単語親密度を予測することを検討することとした。

2 教育基本語彙の分類

小学生の語彙体系としては、基礎語彙、学習基本語彙、学習語彙、一般語彙の 4 分類が提案されている [7]。各語彙体系は相互に図 1 に示すような包含関係を持っている。

基礎語彙は 1,000 語から 2,000 語で言語生活上欠かせないとされている。

学習基本語彙は約 5,000 語で、小学生が十分に駆使することができる語彙とされている。具体的に使用可能な学習基本語彙の 1 つとしては、『語彙指導の方法 [語彙表編]』 [7] に収録されている「学習基本語彙表」(約 4,000 語) がある。

学習語彙は、小学生用の国語辞典に採録されている語彙、約 25,000 語である。この語彙数は、「新教育基本語彙」 [6] などを根拠としている。「新教育基本語彙」には、A(2 段階)、B(3 段階)、C(4 段階) で表現された 9 段階の難易度が単語ごとに付与されている。この学習語彙が小学生の理解できる語彙の上限であるとされている。

一般語彙は、中学生以上を対象にした語彙だが、小学校高学年向けの教科書に出現することもあるため、学習語彙と一般語彙の境界は不明確である。

3 小学生を対象とした単語親密度調査の概要

比較に用いる児童の単語親密度は、以前、単語の小学校教科書における初出学年と親密度の関連を明

らかにした際に調査したデータ [4] を用いた。調査の概要を以下に示す。

3.1 調査対象者

国立大学の附属小学校 2 校の、5 年生 2 クラスずつを調査対象とした。調査対象者数は 2 校合わせて 156 人であった。回答者の性別の比はおよそ 1 : 1 であった。調査対象者が小学 5 年生の学習を一通り完了した時点で実験を行うため、実験時期は 3 月とした。

3.2 呈示刺激

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」[7] から抽出した。抽出された単語から同音異字のある単語や小学校で学習する音読み・訓読み以外の読み方を単語は排除する。最終的に呈示語数は 298 語となった。

3.3 調査方法

調査は調査対象校の教室で行った。スクリーンリーダーの詳細読み策定用の資料とするため、刺激は音声で呈示することとした。男性アナウンサに問題番号に続けて、調査用単語を読み上げさせ、これを音声刺激とした。問題の呈示間隔は 2.5 秒とした。調査の趣旨の説明と回答方法の教示も同じアナウンサの声で収録した。これは、問題が始まる前に読み上げられる。

実験では、カセットテープまたは 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) で再生し刺激呈示を行った。調査の趣旨説明部分を聞かせながら、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節した。

3.4 親密度の評定方法

天野らの成人を対象とした単語親密度調査では、親密度を 1:(低) から 7:(高) の 7 段階で被験者に回答させている [5]。一方、文部省が行った「児童・生徒の語い力調査」では、「よく知っていることば」「だいたいわかることば」「ほ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 [6]。

表 1: 児童の音声単語親密度調査結果 (298 語)

	児童
平均値	80.8
標準偏差	17.9
中央値	87.2
最大値	100.0
最小値	16.7

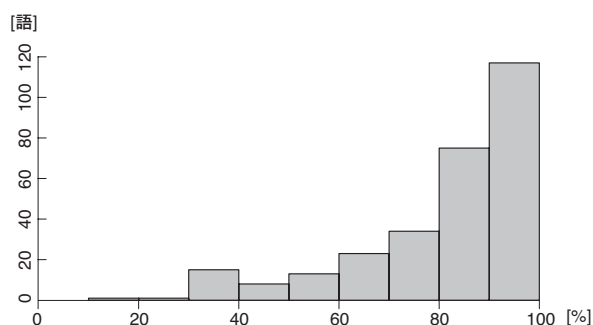


図 2: 児童を対象に調査を行った語の音声単語親密度分布 (298 語)

これらの先行調査を参考に今回の親密度調査の評定を検討した。まず、7 段階の評定尺度法で児童に回答させることは難しいと考えた。次に、4 段階では「だいたいわかることば」と「ほんやりわかることば」の区別がつきにくいと考えた。最終的に今回の親密度調査の評定は次に示す 3 段階で回答させることとした。

ア: 「よく知っている」 知っていると自信を持って言える言葉、自分でも使っている言葉

イ: 「だいたい知っている」 聞いたことはあるけれど、自分ではあまり使わない言葉

ウ: 「知らない」 聞いたこともない言葉

ここでは、各単語について「ア:よく知っている」と答えた者の割合を親密度と定義する。

3.5 調査結果

調査結果の基礎的なデータを表 1 に示す。児童の単語親密度の度数分布を図 2 に示す。

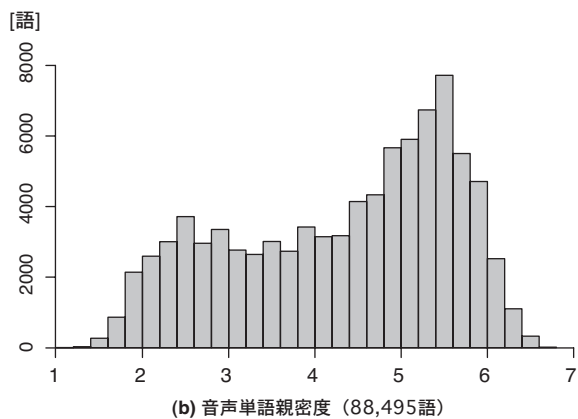
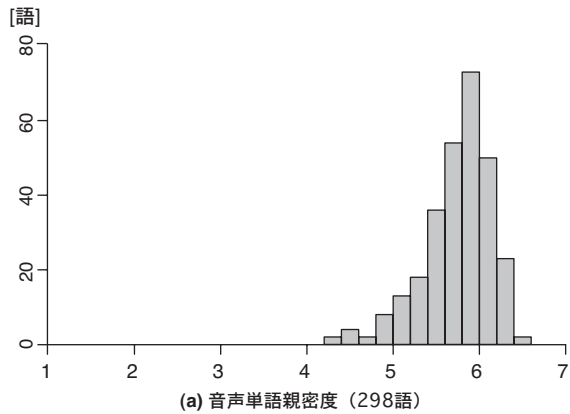


図 3: 音声単語親密度の度数分布. (a) は児童を対象に調査を行った語の分布, (b) は『日本語の語彙特性』 [5] 全体分布

児童の単語親密度の平均値は 80.8 となっている. 分布は親密度が高い方に偏っている. 最大値が 100.0 で最小値が 16.7 と, 広範囲にわたって分布している点が特徴である.

4 単語親密度の比較

前章で述べた調査によって得られた児童の単語親密度と成人の音声単語親密度を比較する.

4.1 成人の音声単語親密度

成人の単語親密度として, 『日本語の語彙特性』 [5] の音声単語親密度を使用した. この音声単語親密度は 32 名または 35 名の被験者が, 呈示された各単語を 1 から 7 の 7 段階で評価したものである.

図 3(a) に児童対象の調査で用いた単語 (298 語)

表 2: 成人の音声単語親密度の分布

	調査対象語 (298 語)	全体 (88,495 語)
平均値	5.740	4.319
標準偏差	0.419	1.270
中央値	5.812	4.625
最大値	6.594	6.719
最小値	4.250	1.029

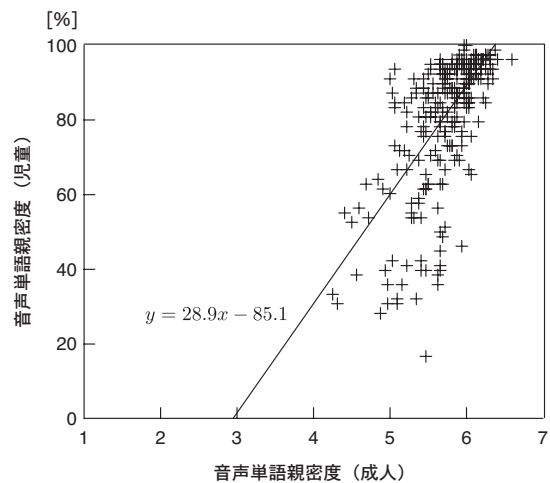


図 4: 児童の単語親密度と成人の単語親密度の関係

の音声単語親密度 (成人) の度数分布を示す. (b) に『日本語の語彙特性』 [5] 全体 (88,495 語) の音声単語親密度の度数分布を示す. 表 2 に成人の音声単語親密度の分布を示す.

図 3(a) と (b) との比較から, 児童対象の調査に用いた単語の分布は, 成人の音声単語親密度において高い方に偏っていることがわかる. 児童を対象とした調査で使用された単語の成人における音声単語親密度の平均値は, 5.740 で全体の音声単語親密度の平均値 4.319 より高い値となっている.

4.2 親密度の関係

x 軸に成人の単語親密度, y 軸に児童の単語親密度をおいた散布図を図 4 に示す.

相関係数は 0.66 で回帰直線は $y = 28.9x - 85.1$ であった.

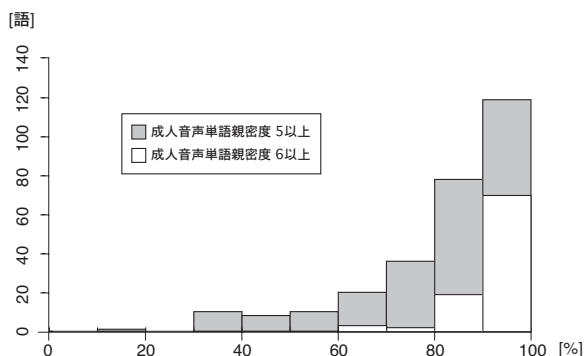


図 5: 成人単語親密度による児童単語親密度の分布の違い

5 考察

5.1 成人の親密度と児童の親密度の関係

児童の音声単語親密度と成人の音声単語親密度との間には、比較的高い相関 ($r = 0.66$) が見られた。このことから、成人の音声単語親密度を基に児童の音声単語親密度を予測できる可能性が示唆されたと考える。

次に成人の単語親密度による児童単語親密度の分布の違いについて考察する。成人の単語親密度が5以上の場合と6以上の場合における児童単語親密度の分布を図5に示す。

成人の音声単語親密度が6以上の場合、児童の音声単語親密度の分布は概ね80%以上に収まっている。児童を対象とした詳細読み作成を行う際、成人の音声単語親密度が6以上の単語を使用するという指標は有効であると思われる。

成人の音声単語親密度が5以上の場合、児童の単語親密度の分布は高い方に偏っているものの、先に挙げた単語親密度6以上の場合に比べ広がっている。つまり、成人の音声単語親密度は比較的高かった語の中でも、児童の音声単語親密度が低かった単語が存在しているということである。その1例を挙げると「批評」という単語では、単語親密度がそれぞれ、成人:5.469、児童:16.7であった。児童を対象とした詳細読み作成を行う際、成人の音声単語親密度が5以上の単語を使用するという指標は、ある程度有効であると思われる。しかし、その指標のみに拠ると児童の単語親密度が低い語を選択してしまう可能性があるため、実際に詳細読みを作成した後、児童に対する分かりやすさの調査を行って、確認す

る必要があるだろう。

5.2 課題

今回の比較に用いた児童の単語親密度は、図3の(a)に示すように、成人の音声単語親密度が高い語を調査時の呈示刺激として用いている。これは、学習基本語彙から調査の単語を選択したことが、原因であると考えられる。成人の音声単語親密度が低い語に関する児童の音声単語親密度を明らかにするためには、学習基本語彙以外からも呈示刺激を作成し調査を行う必要がある。

6 まとめ

小学生を対象とした音声単語親密度調査の結果と成人の音声単語親密度を比較した。両者の間には比較的高い相関が見られ、成人の音声単語親密度から児童の音声単語親密度を予測できる可能性が示唆された。

本研究はスクリーンリーダーの詳細読み策定における資料収集として行っているが、この結果は、子どもを対象とした音声コンテンツ全般に応用できるものとする。

参考文献

- [1] 藤沼輝好, 渡辺恵理子, 鈴木沙耶: スクリーンリーダー使用者のための単漢字詳細説明読みガイドライン, 第27回感覚代行シンポジウム, pp. 67-71 (2001).
- [2] 渡辺文治: 詳細読みについて, *Pin*, No. 23, pp. 32-49 (2002).
- [3] 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—, *信学論*, Vol. J88-D-1, pp. 891-899 (2005).
- [4] 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—, *信学技報*, WIT2005-04 (2005).
- [5] 天野成昭, 近藤公久: 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).
- [6] 国立国語研究所 (編): 教育基本語彙の基本的研究, 明治書院, 東京 (2001).
- [7] 甲斐陸朗: 語彙指導の方法 [語彙表編], 光村出版, 東京 (2002).