

5. 常用漢字・JIS 第 1 水準漢字の詳細読みの策定

Development of *Shosaiyomi* for *Joyo Kanji* and

JIS First Level Kanji

あらし 小学5年配当漢字の詳細読みを新たに作成し、その評価として漢字書き取り実験を行ったところ、既存のスクリーンリーダー製品の詳細読みよりも高い正答率を得ることができた。そこで、教育漢字と同様に利用者の語彙と単語親密度を考慮して、常用漢字（教育漢字を除く）、及び JIS 第 1 水準漢字（常用漢字を除く）の詳細読みを作成することとした。

1. 策定基準

教育漢字に引き続き、常用漢字群と JIS 第 1 水準漢字群の詳細読みも策定することとした。策定基準は教育漢字における基準（第 3 章 2. 1）とほとんど同様だが、(2)初出学年による順位付けは適用せず、(6)詳細読み想起実験の正答率はデータがないため適用できない。また、基準(5)の読みの学習年度は以下の基準に置き換える。

(5)' 常用漢字については、常用漢字表の読みを優先的に使う。

2. 策定作業の進捗

策定手順も教育漢字と同じだが、漢字の難度が高まったため、漢字字典（『漢字源』）と国語辞典（『広辞苑』）にあたって漢字の読み・字体・用例を確認する頻度が高くなった。常用漢字群の詳細読み策定速度は 1 日あたり 140 字から 200 字程度で、939 字を作り終えるには合計 5.5 日間かかった（1 日 6 時間から 8 時間の作業）。JIS 第 1 水準漢字群ではこれより速度が遅く、1 日当たり 90～160 字である。1020 字を策定するのに合計 7.5 日間を要した（合計 13 日間）。

策定作業に携わったのは渡辺哲也と渡辺文治が全作業日の 13 日。研究協力者の岡田伸一が 7 日、同じく研究協力者の吉野嘉那子が 7 日、研究分担者の青木成美が 5 日であった。作業は主として国立特殊教育総合研究所のリエゾンオフィスで行ってきた。このオフィスが JR 田町駅から徒歩 1 分の距離にあることから、そこで作成された詳細読みを「田町読み」と呼ぶことにした。

3. 策定した詳細読みの特性

ここでは、新たに策定した詳細読みの特性を、第 1 章同様に、教育漢字群 1006 字、常用漢字群 939 字、JIS 第 1 水準漢字群 1020 字の 3 群に分けて概説する。

3. 1 単語親密度の分布

所属する単語群による文字音声単語親密度（『日本語の語彙特性』（三省堂，2003））の分布を図 1 に示す。図 1 は箱ひげ図になっており、ひげの上端と下端は最大値と最小値、箱の上端と下端は第 3 及び第 1 四分位数、箱の中央の横棒は中央値を表す。教育漢字群の田町読みの単語親密度は最高値が 6.844、第 3 四分位が 6.281、中央値が 6.062、第 1 四分位数が 5.782、最小値が 4.094 であった。スクリーンリーダ製品の最低値がいずれも 2 以下であることと比べると（第 1 章図 1）、最低値が高く留まっている、つま

り親密度が低い単語を使用していない点が、教育漢字の田町読みの大きな特徴である。

常用漢字群になると単語親密度は全体的に下がるが、それでも中央値が 5.531、第 1 四分位数は 5.094 であり、5 以上を保っている。JIS 第 1 水準漢字群では単語親密度は大幅に下がり、中央値が 5.062、第 1 四分位数が 3.860 となった。これらの数値はスクリーンリーダ製品群と概ね同じである。

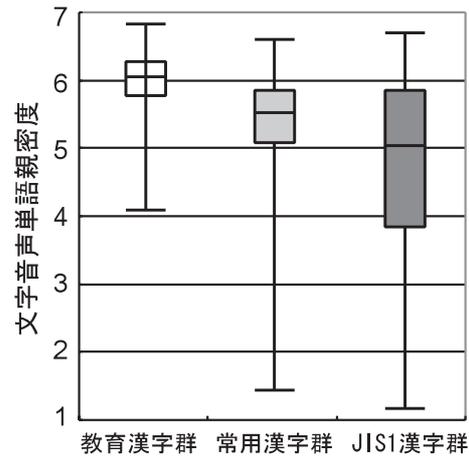


図 1 田町読みの単語親密度の分布

3. 2 詳細読みの構成の割合

所属する単語群によって詳細読みの構成の割合が変化する様子を図 2 に横棒グラフで示す。教育漢字群と常用漢字群では構成の割合はほぼ同じだが、JIS 第 1 水準群になると、大きな変化が見られた。教育漢字群と常用漢字群で最多数を占めた「音を含む熟語の音」と「音を含む熟語の音、訓」の割合が半分程度まで下がった。その分増えたのは、「訓、その他」、「訓を含む熟語の訓」、「音、(その他)」、「音を含む熟語、その他」、「性質の説明の訓」、「性質の説明、その他」「漢字の字形、その他」、「人名・地名」である。JIS 第 1 水準群で構成の割合の種類が多様化する状態はスクリーンリーダ製品の詳細読みと同様な特性となった。なお、構成の分類方法について詳しくは第 1 章の表 1 を参照されたい。

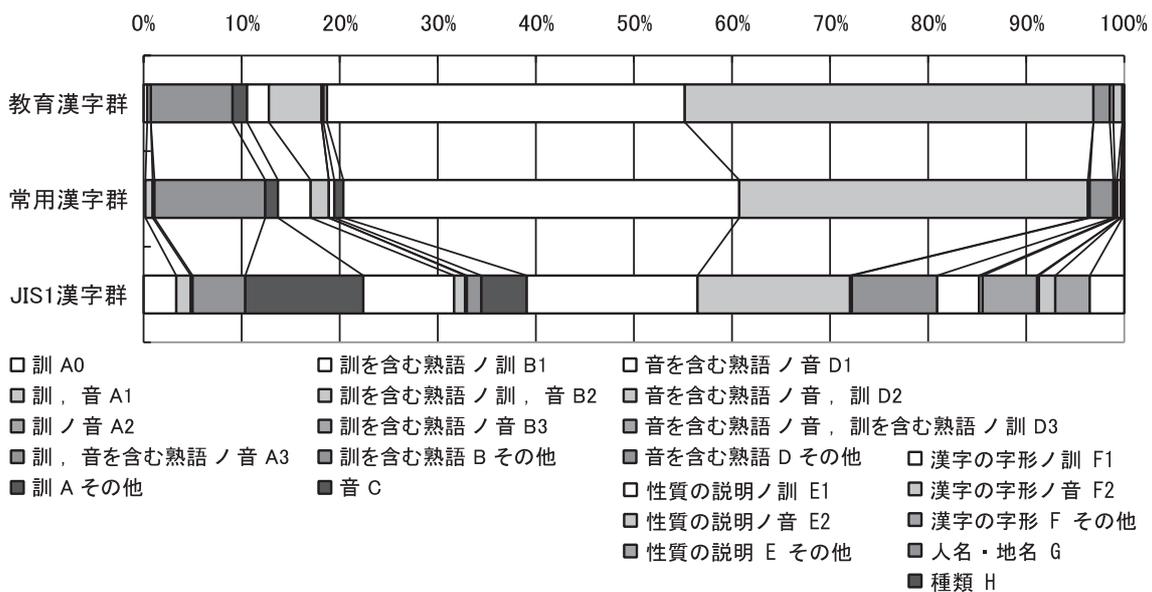


図 2 田町読みの構成

3. 3 スクリーンリーダー製品との一致の度合い

スクリーンリーダー製品の詳細読みと田町読みの一致／類似度を3群に分けて見たのが表1(a)～(c)である。5製品の中ではPC-Talkerとの完全一致率が最も高く、教育漢字群と常用漢字群で30%前後である。説明に用いた単語は同じだが順序が異なるものは多く、JAWS 6.2とは全漢字群にわたって約5割がこれに相当した。どのスクリーンリーダーの詳細読みとも完全一致しない読み表現の数は、教育漢字群では492個(48.9%)、常用漢字群では488個(52.0%)、JIS第1水準漢字群では592個(58.0%)であった。

表1 スクリーンリーダー製品の詳細読みと田町読みの一致／類似度

(a) 教育漢字群

SR	完全一致数 (割合)	部分一致数 (割合)	順序違い数 (割合)
PC-Talker	284(28.2%)	3(0.3%)	30(3.0%)
95Reader	117(11.6%)	0(0%)	33(3.3%)
VDM	149(14.8%)	0(0%)	44(4.4%)
JAWS 4.5	141(14.0%)	1(0.1%)	0(0%)
JAWS 6.2	100(9.9%)	0(0%)	570(56.7%)

(b) 常用漢字群

SR	完全一致数 (割合)	部分一致数 (割合)	順序違い数 (割合)
PC-Talker	292(31.1%)	45(4.8%)	346(36.8%)
95Reader	125(13.3%)	46(4.9%)	468(49.8%)
VDM	135(14.4%)	23(2.4%)	430(45.8%)
JAWS 4.5	169(18.0%)	0(0%)	363(38.7%)
JAWS 6.2	72(7.7%)	0(0%)	538(57.3%)

(c) JIS 第1水準漢字群

SR	完全一致数 (割合)	部分一致数 (割合)	順序違い数 (割合)
PC-Talker	228(22.4%)	24(2.4%)	424(41.6%)
95Reader	172(16.9%)	28(2.7%)	437(42.8%)
VDM	137(13.4%)	14(1.4%)	504(49.4%)
JAWS 4.5	133(13.0%)	0(0%)	388(38.0%)
JAWS 6.2	93(9.1%)	0(0%)	515(50.5%)

4. 策定に困難を伴った漢字

詳細読みの策定に手こずった漢字をいくつか例示する。

- (1) その漢字を含む単語がほとんどないため、親密度が低い単語を使わなければならないもの：「璽」「ギョメイギョジノジ」（天皇の印）「御璽」の単語親密度は 2.0 と低い。
- (2) 常用漢字で代用することが多く、用例・読み・意味も同じ漢字：「稀」（常用漢字は「希」）「マレ、ノギヘンニキボウノキ」字形で説明するほかない。
- (3) 異体字がある漢字：「鈎」「カギ、カネヘンニツツミガマエニム」異体字「鈎」があり、どちらも「ク」、「コウ」「かぎ」と読み、意味も同じなため、字形で説明するほかない。
- (4) 異体字ではないが、意味と読みが一致している字があるもの：「郭」と「廓」はいずれも「カク」「くるわ」と読み、意味もほとんど同じ。用例が異なる熟語で読み分けることとし、「郭」を「リンカクノカク、クルワ」、「廓」を「ユウカクノカク、クルワ」とした。
- (5) 名詞の用例がない漢字：「且」「ナオカツノカツ」
- (6) 地名でしか使わない漢字：「曾」「キソジノソ」、「櫃」「カシハラジングウノカシ」
- (7) 人名でしか使わない（と思われる）漢字：「馨」「ジンメイノカオル、ケイ、カグワシイ」
- (8) 中国人の名前でしか使わない漢字：「鄭」「ジンメイノテイセイコウノテイ」鄭成功：明末の遺臣。鄭芝竜の長子。名は森（しん）。母は肥前平戸の人，田川氏。清軍と戦い，廈門（アモイ）・金門島，のち台湾に拠って奮戦したが，志を遂げず病没。明室から国姓朱氏を賜り国姓爺（こくせんや）と称。近松門左衛門作「国姓爺合戦」により「和藤内（和唐内）」の名で親しまれる。（1624～1662）（広辞苑第五版）
- (9) 策定者らも初めて見る漢字：「蔚」「クサカンムリニリクグンタイイノイ」「蔚然」（心のむすぼれて晴れないさま）という単語があるが、『新明解国語辞典』の見出し語にはなっていない。

出 典

本章は、以下の掲載原稿をもとに再構成した。

- 渡辺哲也, 渡辺文治, 岡田伸一: スクリーンリーダーの詳細読み, その3～単語親密度に基づいた詳細読みの作成 — 田町読み~, Pin, No.27, pp.37-46, 視覚障害情報機器アクセスサポート協会, July 2006.