１．スクリーンリーダ製品の漢字詳細読みの分析

― JIS 第１水準漢字を対象に ―

Variation in the Construction and Word Familiarity of Current *Shosaiyomi* with the Change of Difficulty of Kanji Characters

あらまし　スクリーンリーダ製品の漢字詳細読みで用いられている単語の親密度，及び説明方式（詳細読みの構成）を分析した。JIS 第 1 水準漢字 2965 字を，教育漢字群 1006 字，常用漢字群 939 字，JIS 第 1 水準漢字群 1020 字という 3 段階の難易度に分けて親密度と構成を比較したところ，難易度が上がるにつれて単語親密度が下がること，漢字の読み（音読み／訓読み）と熟語のみで説明できる漢字の数が減り，その分，字形・地名・人名による説明が増えることなどを明らかにした。

１．分析方法

１．１　分析対象

5 種の Windows 用スクリーンリーダ（PC-Talker XP, 95Reader 6.0, VDMW300-PC-Talker, JAWS 4.5, JAWS 6.2）について，JIS 第 1 水準の漢字 2965 字の詳細読みを書き起こし，これを分析対象とした。

１．２　分析方法

１．２．１　説明単語の抽出と親密度の算出

詳細読みに用いられている単語を抽出した。詳細読みの構成要素のうち訓読みはそれ単独で意味が通るため単語として扱った。他方で音読みは，それ単独で単語となる場合が少ないため，単語として抽出しなかった。性質による説明の中に現れる単語（例えば，「～を意味する」），及び字形による説明の中に現れる単語（偏や旁の名称，位置を表す語「上」「下」など）も抽出しないこととした。

抽出した単語は NTT のデータベース[1]にあたって親密度の数値を求めた。このデータベースで調べられるのは『新明解国語辞典第四版』（三省堂，1989）の見出し語に限られる。単語親密度には，単語を音声で聞いた場合の音声単語親密度，文字で見た場合の文字単語親密度，文字と音声の両方を提示された文字音声単語親密度の 3 種類がある。音声で提示される詳細読みに適用するには音声単語親密度をみるのが妥当だが，日本語で 3 割から 4 割にも上る同音異義語を音声単語親密度では区別できないため，今回は文字単語親密度を親密度データとして用いる。単語親密度は，1 から 7 の間の数値で示され，数値が高いほど親密度が高い。

１．２．２　詳細読みの構成の分類

詳細読みをその構成によって分類した。詳細読みの構成とは，説明に使う要素（音読み，訓読み，当該漢字を含む単語，漢字の意味を表す句表現など）とその出現順序の組み合わせである。全ての組み合わせ数は約 80 種類に及んだが，ここでは表 1 のように 22 種類にまとめ，各構成の割合を見る。

まず，当該漢字の訓読みが最初に現れるものを A グループとした。その小分類として，訓読みの後に読点と音読みをつなげた A1，訓読みと音読みを「ノ」で接続した A2，第 2・第 3 要素として「音読みを含む熟語ノ音読み」が来る A3 を設けた。第 2 要素としてこのほかに，対象の性質や漢字の字形，異なる訓読みなどがあるが，数が少ないので，表では「その他」にまとめた。

表 1　詳細読みの構成の分類の定義と例

分類記号	構成	詳細読みの例（括弧内は元の漢字）
A0	訓	サラ(皿)
A1	訓，音	マナブ，ガク(学)
A2	訓ノ音	ツカウノシ(使)
A3	訓，音を含む熟語ノ音	ヒ，カヨウビノカ(火)
A その他	訓，＊	ヒロ，ユミヘンノコウ(弘)
B1	訓を含む熟語ノ訓	カイガラノカイ(貝)
B2	訓を含む熟語ノ訓，音	オオアメノアメ，ウ(雨)
B3	訓を含む熟語ノ音	カワギシノセン(川)
B その他	訓を含む熟語，＊	ニモツノニ，シュッカノカ(荷)
C	音，＊	エン，ドウブツノサル(猿)
D1	音を含む熟語ノ音	バンサンノサン(餐)
D2	音を含む熟語ノ音，訓	ヒゲキノヒ，カナシイ(悲)
D3	音を含む熟語ノ音，訓を含む熟語ノ訓	シソンノシ，コドモノコ(子)
D その他	音を含む熟語，＊	カイフクノカイ，リッシンベンニハイ(恢)
E1	性質の説明ノ訓	クサキノキリ(桐)
E2	性質の説明ノ音	カンスウジノイチ(一)
E その他	性質の説明	タベルアメ(飴)
F1	漢字の字形ノ訓	キヘンニフユンヒイラギ(柊)
F2	漢字の字形ノ音	サンズイニスエノマツ(沫)
F その他	漢字の字形，＊	キヘンノモリ，トウジノトウ(杜)
G	人名・地名	ナハシノナ(那)
H	種類，＊	イタイジノクノク(駈)

次に，訓読みを含む熟語が第 1 要素となるものを B グループとした。訓読みを含む熟語と訓読みが「ノ」で接続された表現を B1 とし，B1 の後ろに読点と音読みが来るものを B2 とした。更に，訓読みを含む熟語と音読みが「ノ」で結ばれた表現を B3 とし，これ以外をその他とした。

C グループは，音読み単独が最初に現れるものである。第 2 要素は訓読みを含む熟語

が最も多いが，このグループは相対的に数が少ないので小分類はしない。

D グループは，当該漢字の音読みを含む熟語が最初に来るもので，そのうち全スクリーンリーダに共通して多い構成は，熟語に続けて「ノ」及び音読みが来る D1 である。説明の熟語としては 2 字熟語が多く，これは同音異字のものが多い。そこで区別のため，D1 の後ろに読点と訓読みを追加した表現 D2，D1 の後ろに「訓読みを含む熟語ノ訓読み」を追加した表現 D3 などがある。

E，F，G，H グループはそれぞれ，漢字の意味による説明，漢字の字形による説明，当該漢字を含む人名・地名が最初に来る表現，漢字の種類による説明である。

１．２．３　漢字群の分類

JIS 第 1 水準 2965 字を，教育漢字 1006 字，常用漢字から教育漢字を除いた 939 字，JIS 第 1 水準漢字から常用漢字を除いた 1020 字の 3 群に分け，それぞれを教育漢字群，常用漢字群，JIS 第 1 水準漢字群と呼ぶことにする。この順序で漢字の難易度が上がると見なす。分析した単語親密度と構成の割合について，スクリーンリーダ製品ごとに漢字の難易度による変化を見る。

２．分析結果

２．１　単語親密度分布の変化

図 1 に，単語親密度分布の漢字難易度による変化をスクリーンリーダごとに示した。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。いずれのスクリーンリーダにおいても，教育漢字群の詳細読みの単語親密度の中央値は 6 前後，第 1 四分位数も 5.6 前後と高い値であった。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 5.5 前後，第 1 四分位数は 5 前後を保っている。単語親密度 5 以上の単語は成人の 95%くらいの人が知っているとされるので[1]，常用漢字までの詳細読みの説明語の約 75%には，ほとんどの成人が知っているとされる単語が使われていると言える。

JIS 第 1 水準漢字群では単語親密度は大幅に下がり，スクリーンリーダ 4 製品で中央値が 5 前後，第 1 四分位数が 3.7～3.9 となった。1 製品はこれより低く，中央値が 4.7，第 1 四分位数が 3.4 であった。単語親密度 4 の単語を知っている成人の割合は 50%程度なので，説明語を知っている人の割合が半分以下の単語が詳細読みの 4 分の 1 を占める計算となる。

最小値に注目すると，教育漢字群，及び常用漢字群の中にも単語親密度が 1.5 から 2 となるものが含まれていた。単語親密度 2 はほぼ誰も知らない単語であり，数は少なくてもこのような単語が含まれていたことが，特に児童にとっては詳細読みからもとの漢字を想起する際の問題となっていたことが推察される[2]。

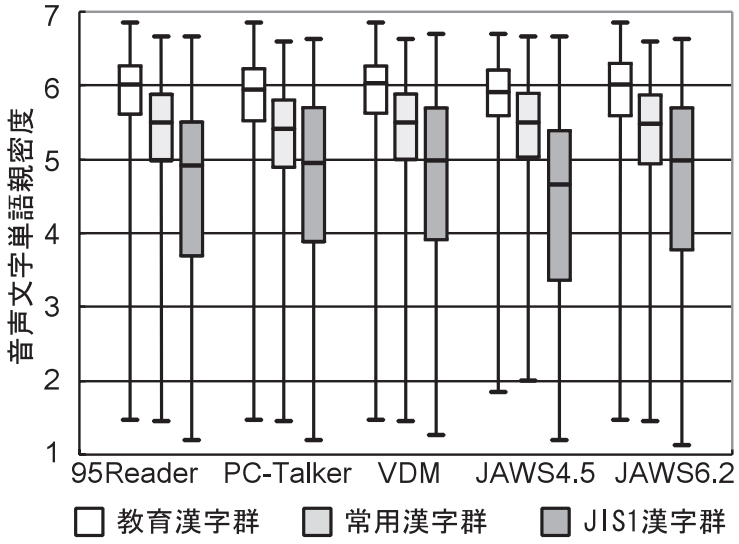

図 1　単語親密度の漢字難易度による変化　箱ひげ図のひげの上端と下端は最大値と最小値，箱の上端と下端は第 3 及び第 1 四分位数，箱の中央の横棒は中央値を表す。各スクリーンリーダの三つのグラフのうち，左側白い箱は教育漢字群，中央薄いグレーの箱は常用漢字群，右側濃いグレーの箱は JIS 第 1 水準漢字群を表す。

２．２　詳細読みの構成の割合の変化

詳細読みの構成の割合が漢字難易度に従って変化する様子を図 2 に横棒グラフで示す。構成の割合にはスクリーンリーダ製品ごとの特徴が見られる。教育漢字群を見ると，95Reader では，訓読みが最初に来る A グループが 3 分の 2 を占め，残る 3 分の 1 は D1 の「音読みを含む熟語ノ音読み」であり，短い表現が主体となっている。

PC-Talker と JAWS の詳細読みは，音読みを含む熟語が最初に来る D グループに殆ど含まれる。PC-Talker で最も多いのは，音読みの次に訓読みを示す D2 で 72.3%，JAWS では，音読みのみの D1 が全体の 89.6%を占めた。VDM は A と D の 2 グループが半分弱ずつを占める。小分類で最も多いのは D1 である。A グループの中では，「訓読みノ音読み」となる A2 が 26.0%と多いのが他製品と異なる。いずれのスクリーンリーダにおいても，性質や字形による説明が最初に来る E 及び F グループの割合は 3%以下と低かった。

常用漢字群となり漢字の難易度が高くなると，各構成の割合も変化した。95Reader では，訓読みの後に音読みを含む熟語のみ(A3)，あるいは音読みを含む熟語と訓読みという構成(D2)が増えていった。PC-Talker では，訓を含む熟語の割合が上がった。JAWS 6.2 では音読みを含む熟語(D1)の割合が高くなり，その分 A3 の割合が減った。VDM と JAWS 4.5 では，教育漢字群と常用漢字群の間で割合の変化は見られなかった。

難易度の低い漢字では音読み，訓読み，熟語を使った構成がいずれのスクリーンリーダでも頻用されたが，JIS 第 1 水準漢字群ではこれらの割合が減り，その代わり漢字の意味による説明，漢字の字形による説明，人名・地名による説明が増えた。性質や字形などは，スクリーンリーダ製品によっては 15～20%の割合に達した。また，用いられている構成の種類が増えたことも特徴的である。漢字の難易度が上がるにつれ，一定の方式だけでは説明しきれないという詳細読み策定の難しさがうかがえる。

3．考　察

漢字は，字形，読み，意味の三つの要素を有する。視覚障害者に漢字を判別させるには，視覚情報である字形を除いて，読みと意味，そして意味から派生した熟語や用例を使うのが適切である。既存スクリーンリーダの詳細読みを分析したところ，常用漢字ま

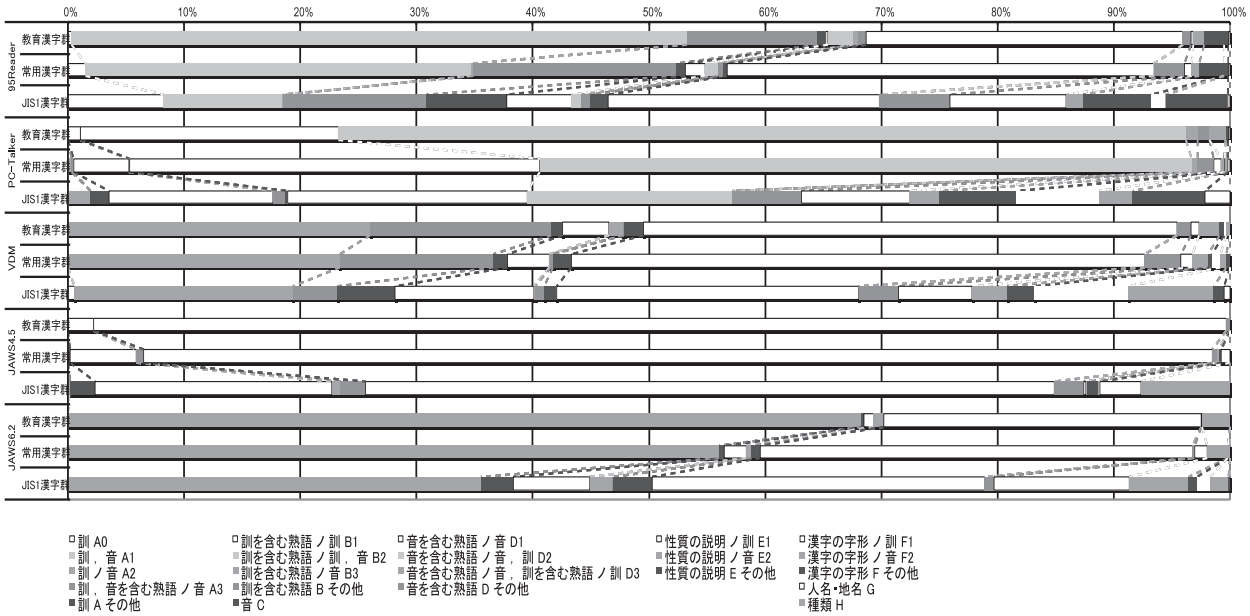

図 2　詳細読みの構成の割合が漢字難易度に従って変化する様子　要素数が多いため，同じ色が異なった要素に用いられている。構成分類の大きなまとまり（A，B，D，E，F）ごとに白色から始まることを参考に各要素を同定されたい。この図の読み取り方については本文を参照してほしい。

では読みと熟語・用例で大部分の漢字の説明ができていた。しかし，常用漢字以外の JIS 第 1 水準漢字となると，字形や漢字の種類（新／旧字，異体字等）で説明することが避けられない場合があった。それらには以下の 2 種類がある。

- 異体字があるもの。異体字同士は読みと意味が同じなため，字形の違いでしか説明できない。例:「鈎」（カギバリノカギ）と「鉤」「カネヘンニハイクノク」(PC-Talker)。
- 常用漢字外のため，常用漢字で代用することが多く，読み・意味・用例とも，対応する常用漢字と一致するもの。例 :「稀」（常用漢字外）と「希」（常用漢字）。「廓」（常用漢字外）と「郭」（常用漢字）。

その漢字の条件から字形で説明するのもやむを得ないとしたが，漢字を見た経験のない視覚障害者にとって字形による説明は一般的には理解しづらいので[3]，その使用は最低限にとどめるべきであろう。ただし，学習により視覚障害者の漢字の知識は変化するともされており，今後長期的な研究が必要である。

JIS 第 1 水準漢字でもう一つ特徴的なのは，人名・地名を使った説明表現の増加である。人名・地名以外に用いない漢字では，これらによる説明もやむを得ない。しかし，どの範囲の人名・地名まで使ってよいかという基準が必要であろう。例えば地名であれば県名と県庁所在地までに限るのか，市区町村でもよいのか，人名ならどの資料に載っているものまでか，などの基準である。

これまで明らかにしてきたとおり，詳細読みを聞いて漢字を判別できるかどうかは利用者の知識に大きく依存する[4]。常用漢字であれば，平均的な成人の語彙を前提に詳細読みを策定できるが，常用漢字以外の JIS 第 1 水準漢字となると，比較的高い知識水準の利用者を想定しなければ，たとえどのような説明表現をしても判別は難しいものと思われる。

謝　辞

詳細読みの分析には，千葉大学の学部生／大学院生（当時）の福田慧人氏，大須賀智子氏，矢島直人氏，神谷佐武郎氏，及び当研究所で非常勤勤務している村上真理氏に御協力頂いた。

参考文献

[1] 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.

[2] 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.

[3] 海保博之, 佐々木正人, “先天盲の漢字存在感と漢字検索過程, ” 特殊教育学研究, Vol.21, No.4, pp.7-16, 1984.

[4] 渡辺哲也, 渡辺文治, 山口俊光, 大杉成喜, 澤田真弓, 岡田伸一, “スクリーンリーダの詳細読みによる漢字想起実験―成人を対象とした場合―, ” 電子情報通信学会論文誌 D, Vol.J89-D, No.3, pp.602-605, 2006.

出　典

本章は，以下の原稿をもとに再構成した。

- 渡辺哲也, 山口俊光, 渡辺文治, 岡田伸一 : スクリーンリーダの漢字詳細読みに関する研究―既存詳細読みの構成及び単語親密度の漢字難易度による変化―, 電子情報通信学会技術報告, WIT2006-31, October 2006.
- 渡辺文治, 渡辺哲也, 大杉成喜, 岡田伸一 : 視覚障害者のための詳細読みの検討　その３―教育漢字の詳細読みにおける単語親密度の比較―, 日本特殊教育学会第 44 回大会発表論文集, p.332, 前橋, September, 2006
- 渡辺文治, 渡辺哲也, 大杉成喜, 澤田真弓, 岡田伸一 : 視覚障害者のための詳細読みの検討 その２ ―教育漢字とそれ以外の JIS 第 1 水準漢字の分析―, 日本特殊教育学会第 43 回大会発表論文集, p.596, 金沢, September, 2005.