

国立国語研究所の漢字研究関係資料

高田智和・寺島宏貴・中島彩花（国立国語研究所）

1. はじめに

本発表は、第 37 回「東洋学へのコンピュータ利用」研究セミナーを国立国語研究所で開催するにあたり、休憩時間に実施する展示について、展示内容を紹介するものである。今回の展示内容は、「東洋学へのコンピュータ利用」にちなみ、国立国語研究所が所蔵する漢字と漢字情報処理に関する研究資料から選んだものである。

2. 1948 年読み書き能力調査資料

1948 年、第 2 次世界大戦後の占領下において、1946 年の第 1 次アメリカ教育使節団報告書に基づき、GHQ/SCAP/CIE（連合国軍最高司令官総司令部民間情報教育局）の指示により、教育研修所（国立教育研究所の前身）を中心に読み書き能力調査委員会を設置して、読み書き能力調査が実施された（以下「1948 年調査」）。配給台帳等に基づくランダムサンプリングによって、全国 270 地点、15～64 歳の男女 16,820 人を対象に、全 90 問の読み書き能力テストを行い、その後の読み書き能力調査をはじめとして大規模学力調査や計量的社会調査の出発点となった調査である（木村拓也 2006、佐藤寧 2014）。1948 年調査の日本語報告書は 1951 年 4 月に東京大学出版部から刊行されている（読み書き能力調査委員会 1951、以下「1951 年報告書」）。

国立国語研究所の設立は 1948 年 12 月 20 日であり、1948 年調査の実施直後である。設立当初の職員には、柴田武氏や野元菊雄氏など 1948 年調査に携わった者がおり、その後の国立国語研究所の「科学的な国語研究」に多くの知見をもたらした。ランダムサンプリングによる言語意識の調査は 1948 年調査の手法を言語調査に適用したものであり、雑誌・新聞等の語彙調査も 1948 年調査のテスト作成の経験をふまえたものと見なされよう。

ランダムサンプリングの実施には統計学の専門的知識が必要不可欠である。1948 年調査には統計数理研究所の林知己夫氏や白石一誠氏などが参加し、後に国立国語研究所が企画した鶴岡調査（1950 年）や岡崎敬語調査（1953 年）にも統計数理研究所の職員が参画している。国立国語研究所が所蔵する 1948 年読み書き能力調査資料は、1948 年調査に参加した当時の統計数理研究所の職員の持ち物だったものである。なぜ統計数理研究所ではなく国立国語研究所に伝来しているのか、その理由は明らかではない。しかし、設立当初の国立国語研究所が、統計数理研究所と密接なつながりを持っていたことは、現存資料からも明

らかだと言えるだろう。なお、1948年調査に参加した当時の国立国語研究所の職員の資料の存在は未だ明らかにはなっていない。

“LITERACY RESEARCH PROGRAM” (fo0161-i0010)

“LITERACY RESEARCH PROGRAM”は、1948年調査の企画書と目される文書である。1ページ当たり55行程度のタイプライターによる英文文書で、全12ページ(約6,400語)である。CIE民間情報教育局の文書群にも“LITERACY RESEARCH PROGRAM”が存する(GHQ/SCAP Records, Civil Information and Education Section(民間情報教育局(略称:CIE)文書)ボックス番号5914;フォルダ番号30、アメリカ国立公文書館現蔵、国立国会図書館所等がマイクロ資料所蔵)。目録カードに1948年3月と記載され、アイテムタイトルは「Literacy Research Program Distributed among the Committee Members」であり、これに従えば、1948年3月の読み書き能力調査委員会の委員に配布された資料ということになる。CIE文書の“LITERACY RESEARCH PROGRAM”と国立国語研究所所蔵の“LITERACY RESEARCH PROGRAM”は、行取りと枚数が全く同じであり、「同版」である。管見では、アメリカ国立公文書館と国立国語研究所、あともう一か所スタンフォード大学フーヴァー研究所(CIE民間情報教育局に所属したジョゼフ・トレイナー氏の文書群に存する)のほかに“LITERACY RESEARCH PROGRAM”の現存情報は知られておらず、目下のところ国内に現存する“LITERACY RESEARCH PROGRAM”は国立国語研究所所蔵資料のみである。

なお、国立国語研究所とROIS-DS社会データ構造化センターは、LRPの学術的重要性に鑑み、2022年3月に“LITERACY RESEARCH PROGRAM”をインターネット公開している(<https://dglb01.ninjal.ac.jp/ninjal/bunken.php?title=literacyrpen,CC-by>)。“LITERACY RESEARCH PROGRAM”を読み解くには、言語学、教育学、心理学、社会学、統計学の協業が必要であり、今後の課題である。

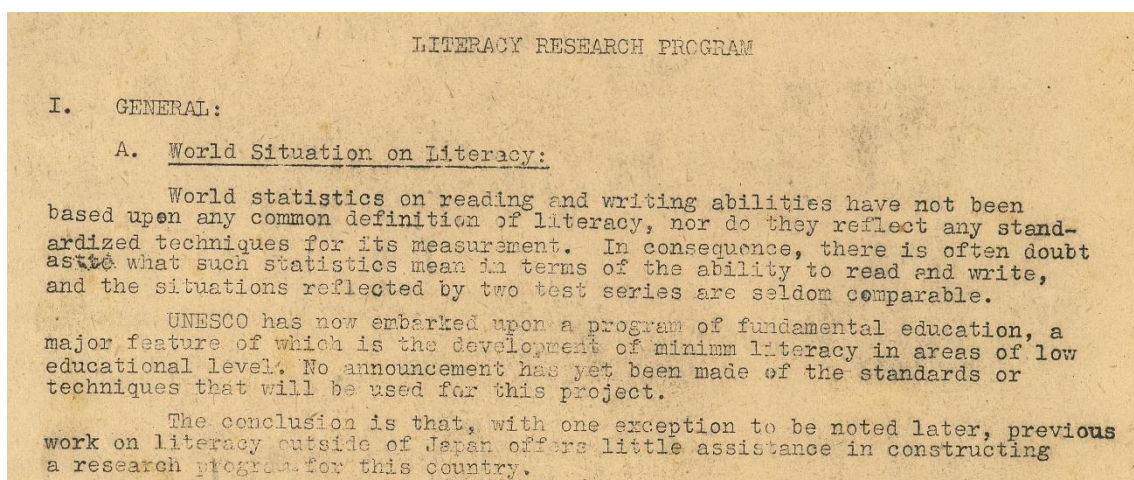


図1: “LITERACY RESEARCH PROGRAM” (冒頭)

「統計日誌」(fo0161-i0003)

「統計日誌」は、1948年調査(1948年3月から1949年3月までの13か月間)のサンプリングや調査結果の集計、報告作成に費やした時間と人員を記録したものである。わずかに1枚のメモ程度のものであるが、1951年報告書の「経過日曆(附. アルバイト使用状況)」(pp.411-414)の草稿と考えられる。1948年調査はIBMの計算機を使って調査結果の集計をしたことが知られているが、「統計日誌」には集計結果の受領日が、1948年12月上旬から1949年3月上旬まで合計5回記されている。「統計日誌」には「6回 1949 ?」と6回目の記述があるが、1951年報告書に6回目の記載がないため、IBMからの集計結果の提供は合計5回であったとみるべきだろう。

IBM 手続日		
1回	1948	12月上旬
2回	1948	12月下旬
3回	1949	2月上旬
4回	1949	2月下旬*
5回	1949	3月上旬 4日
6回	1949	?

IBM システム プログラム

図2: 「統計日誌」(一部)

「Hityôstasy no Kôsei to Sono Heikinten.」(fo0161-i0026)

「Hityôstasy no Kôsei to Sono Heikinten.」はホチキス止めの小冊子で、17ページにわたって1948年調査の集計結果の数表(「SYOKUGYÔ-betu Kôsei(職業別構成)」「SEI-betu Kôsei(性別構成)」など)が記載されたものである。表紙には「Hityôstasy no Kôsei to Sono Heikinten. (TUZUKI)」とあり、また、ノンブルが「24」から始まるため、全体の後半部と考えられるが、前半部は国立国語研究所には残存していない。表紙右上には「102部ノ中ノ第14号」のようにナンバリングがされた配布物である。国立国語研究所には「Hityôstasy no Kôsei to Sono Heikinten.」が34冊ある。

また表紙には「'49n. 7gt.」とあって、1949年7月作成であることがわかる。1951年報告書の800.0一般経過には、「(1949年)7月16日 東京大学東京大学法文経第25番教室で読み書き能力調査報告会を開く。」(p.410)とあり、当日報告会プログラムも記載している。

「Hityôstasy no Kôsei to Sono Heikinten.」は1949年7月16日読み書き能力調査報告会の配布資料の残部ではないかと思われる。

Dai 33 Hyō *Agud* Sei-betu Kōsei

	Sibu			Gunbu				Gōkei				
	OTOKO	ONNA	Humei	GŌKEI	OTOKO	ONNA	Humei	GŌKEI	OTOKO	ONNA	Humei	GŌKEI
HOKKAIDŌ	116	108		224	288	244		532	404	352		756
	51.8	48.2		100.0	54.2	45.8		100.0	456.5	398.0		854.5
TŌHOKU	184	172		356	851	786	3	1640	1035	958	3	1996
	51.7	48.3		100.0	51.9	47.9	0.2	100.0	989.4	915.7	29	1908.0
KANTŌ	949	814		1763	1625	1637	2	3264	2574	2451	2	5027
	53.8	46.2		100.0	49.8	50.7	0.1	100.0	2725.1	2583.9	2.0	5311.0
KANSAI	867	854	1	1722	1187	1249		2436	2054	2103	1	4158
	50.3	49.6	0.1	100.0	48.7	51.3		100.0	2096.5	2145.9	1.0	4243.4
TYŪ-SIKOKU	295	274		569	873	871	2	1746	1168	1145	2	2315
	51.8	48.2		100.0	50.0	49.9	0.1	100.0	1182.1	1158.0	2.0	2342.1
KYŪ-SYŪ	365	383		748	897	923		1820	1262	1306		2568
	48.9	51.2		100.0	49.2	50.8		100.0	1200.1	1241.9		2442.0
ZENKOKU	2776	2605	1	5382	5721	5710	7	11438	8497	8315	8	16820
	51.5	48.5	0.0	100.0	50.0	49.9	0.1	100.0	8649.7	8443.1	7.9	17101.0

図3: 「Hityōsasya no Kōsei to Sono Heikinten.」 (Dai 33 Hyō)

3. 漢字テレタイプ資料

国立国語研究所は語彙調査の大規模化を企図し、1963[昭和 38]年から 1964[昭和 39]年にかけて調査研究課題「大量語彙調査機械化のための準備的研究」において技術的検討を行い、1965[昭和 40]年に大型計算機並びに漢字テレタイプを導入した。漢字テレタイプは沖電気工業製のものを3台購入している。1965年導入の漢字テレタイプは現存していないが、1967[昭和 42]年に追加で導入した漢字テレタイプ（ケン盤セン孔機ケン盤部と漢字印刷機印刷部）を保存している。漢字テレタイプによる漢字処理方法と初期の研究成果は『国立国語研究所報告 31 電子計算機による国語研究』（1968年）にまとめられている。

なお、1968年新興製作所製造の漢字テレタイプ（SCK-201形漢字鍵盤さん孔機）が情報処理技術遺産に登録されている（<https://museum.ipsj.or.jp/heritage/SCK-201.html>）。報道機関で利用が始まった漢字テレタイプは、昭和40年代に言語研究を目的とする言語処理分野に利用域が広がったのである。

「漢テレコード表」(fo0064-i0021)

「漢テレコード表」は、国立国語研究所が導入した漢字テレタイプのコードブックである。当該資料は書入れが多数あることから、コードブックの草稿であると考えられる。コードブックには、部首順や代表五十音順のものも作成されている。盤外字の逆索引もある。なお、コードブックと盤外字逆索引は、国立国語研究所研究図書室蔵書が国立国語研究所学術情報リポジトリにて公開されている (CC by)。

漢テレコードブック (1967年3月) <https://doi.org/10.15084/00002389>

漢テレ盤外字逆引表 (1971年2月) <https://doi.org/10.15084/0002000052>

1 乃 ○ % Y 經 景 0 7 0 :	乙 ○ M 4 芸 横 0 9 3 :
2 久 ○ 7 : 經 都 0 7 0 N	九 ○ I Z 芸 声 0 9 3 /
3 之 ○) T 經 愛 0 7 2 1	乞 ○ S H 芸 政 0 9 3 X
4 乎 → 經 位 0 7 2 A	也 ○ W Q 芸 勢 0 9 3 Z
乏 ○ U S 經 違 0 7 2 E	乱 ○ X X 芸 切 0 9 4 L
5 乘 ○ I 3 經 影 0 7 3 9	7 乳 ○ 井 Q 芸 設 0 9 4 R
9 乘 乘 乘 ○ 經 石 0 7 3)	10 乾 ○ O E 芸 改 0 9 5 C
下 乘 → 經 世 0 7 2)	龜 ○ (L 芸 線 0 9 5 J
下 乘 → 經 都 0 7 0 N	12 亂 乱 乱 ○ 芸 素 0 9 5 R
下 乘 → 經 井 0 7 2)	

図 5: 漢テレコード表 (一部)

4. 林大 JIS 漢字資料

林大氏 (1913[大正 2]年~2004[平成 16]年) は文部省及び国立国語研究所に籍を置き、当用漢字表 (1946 年)、現代かなづかい (1946 年)、当用漢字音訓表 (1948 年)、当用漢字別表 (1948 年)、当用漢字字体表 (1949 年) など戦後の一連の国語施策に携わった研究者である。国立国語研究所の調査研究では、婦人雑誌の語彙調査 (1950-1953 年) や『分類語彙表』(1964 年刊) などを中心となって担当した (甲斐睦郎 2004)。1976 年から 1982 年まで国立国語研究所第 3 代所長を務めた。没後、2006 年に蔵書や研究資料 (ノート類) が国立国語研究所に寄贈されている (fo0176 林大寄贈資料、現在整理中)。

また、林大氏は漢字コードの標準化、特にレパートリーの選定に大きく貢献していたことが知られている (芝野耕司 2002、池田証寿 2004)。林大氏の寄贈資料には、漢字を登載した世界初の文字コード標準である JIS X 0208 (当初は JIS C 6226) の策定に関わる資料も含

まれている。

『情報交換用漢字符号系 JIS C 6226-1978』第1刷

『情報交換用漢字符号系 JIS C 6226-1978』は、いわゆる JIS 漢字規格の第1刷である。林大氏旧蔵の JIS C 6226 規格票第1刷には、林大氏自身の書入れがある。JIS C 6226 規格票第1刷は誤記の多いことが知られている。赤インクの手入れは、規格票の誤記を訂正したものである。また、当用漢字字体との異同や旧字体との照合結果、人名用漢字との突合結果を、黒鉛筆で印をつけている。刷り上がった規格票第1刷を精密に検証していたことがうかがえるが、常用漢字表の制定（1981年）を前にした検討とも考えられる。

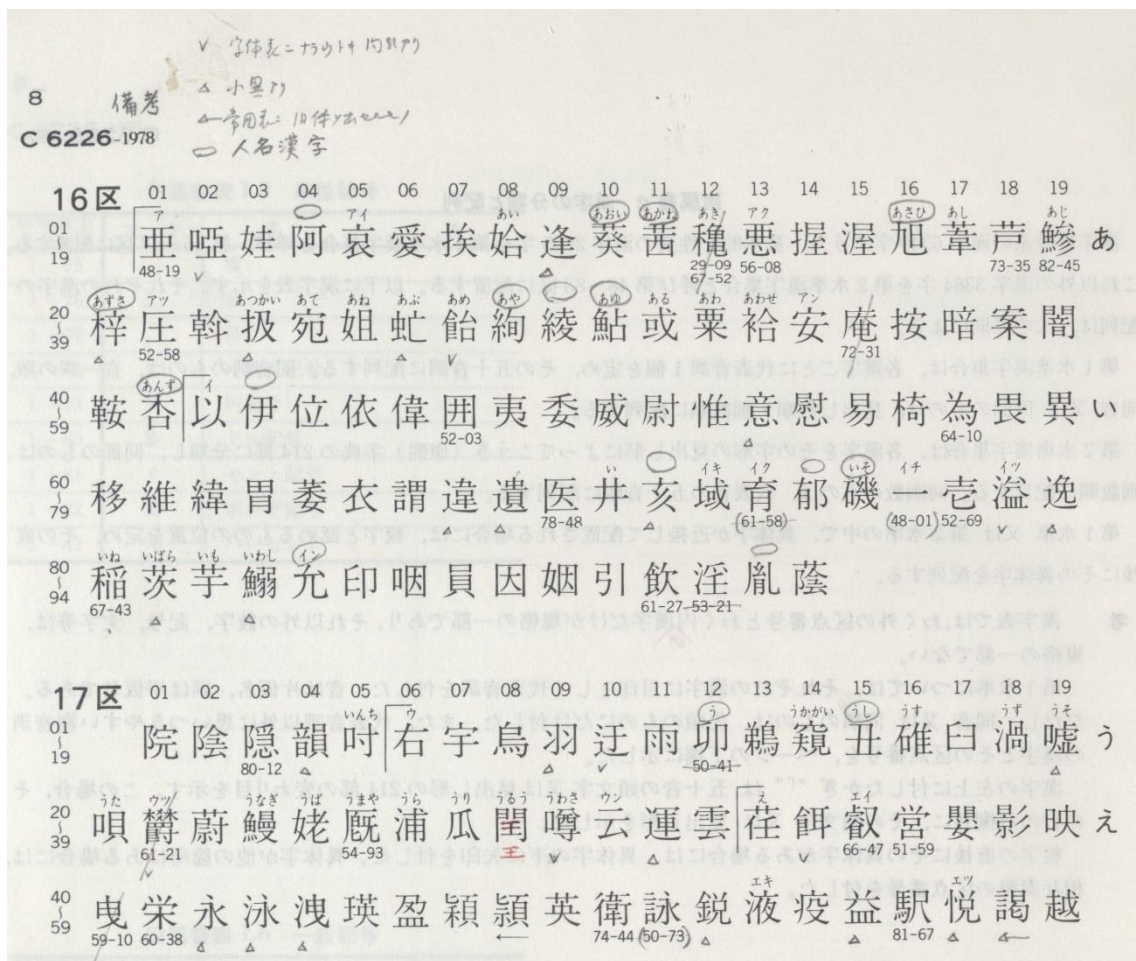


図6: 『情報交換用漢字符号系 JIS C 6226-1978』第1刷（一部）

『標準コード用漢字表（試案）』

『標準コード用漢字表（試案）』は情報処理学会漢字コード委員会が1971年10月に作成したものである。前述の JIS C 6226-1978 の解説には、第1水準漢字及び第2水準漢字の水準分けに使用した37種の漢字表が記載されている。『標準コード用漢字表（試案）』は、そ

ここで第2番目に挙げられた漢字表である。林大氏旧蔵の『標準コード用漢字表（試案）』にも林大氏の自筆書入れ（黒鉛筆）がある。表紙に「（昭和）55年7月」とあり、書入れ時期が明確である。書入れの内容は、当用漢字字体との異同が中心である。常用漢字表の制定（1981年）を前にした検討、あるいはJIS C 6226の見直しに向けた検討であろう。

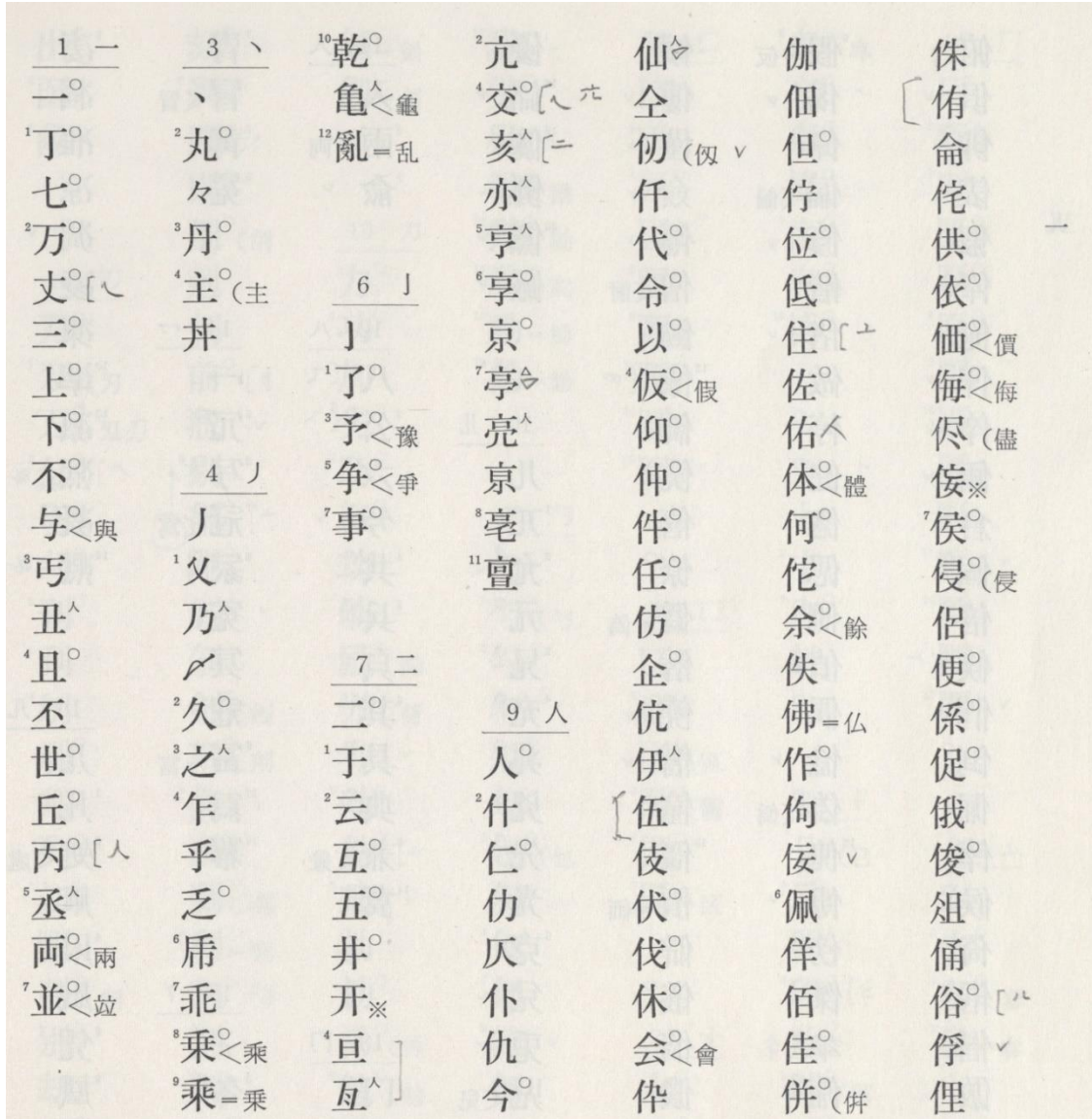


図6: 『情報交換用漢字符号系 JIS C 6226-1978』第1刷（一部）

日下部表

前述の『標準コード用漢字表（試案）』には、13種の資料を用いて漢字選定を行ったことが記されている。その冒頭に挙げられた選定資料が「日下部重太郎「現代国語思潮続編」（昭和8）の附録表」であり、本資料である（以下「日下部表」）。

日下部重太郎（1878[明治9]年～1938[昭和13]年）は、大正から昭和にかけて、朗読法や

国語国字問題に取り組んだ日本語研究者である。「実用漢字」集合の研究も熱心に行い、1920[大正 9]年に『実用漢字の根本研究』（付属表「常用漢字等級表」）、1932[昭和 7]年に『現代国語精説』（付属表「実用漢字等級表」）、1933[昭和 8]年に『現代国語思潮続篇』（付属表「附録 現代日本の実用漢字と別体漢字との調査及び「常用漢字」の価値の研究」）を刊行している。『標準コード用漢字表（試案）』が参照した漢字表は、日下部重太郎氏の最後の3番目の漢字表である（高田智和 2012）。

林大氏旧蔵の日下部表は、『現代国語思潮続篇』付属表のページ見開き複写 18 枚に、赤インクで漢字の周囲に書入れがなされている。書入れは、大西雅雄『日本基本漢字』、『広辞苑』第 2 版付録「通用漢字一覧」、朝日新聞社「統一基準漢字書体表」、全日本漢字排列協議会「常用漢字目録」、日本活字鑄造株式会社「標準活字目録」、国会図書館用 NDL70 用コード表、共同通信社「漢テレハンドブック」、日経 FAM・M 型タイプ文字表の 8 種漢字表との照合結果であり、『標準コード用漢字表（試案）』が挙げる選定資料と一致する。漢字の周りの位置と、記号の形状とで、照合結果を区別して記述する方法は、ヲコト点を連想させる。試みに、人部について『標準コード用漢字表（試案）』と照合してみると、書入れのある漢字はすべて『標準コード用漢字表（試案）』に採用されていることが確認できる。『標準コード用漢字表（試案）』の選定基準（日下部表にある漢字で、大西雅雄『日本基本漢字』など 8 種漢字表のいずれかに現われる漢字を採用する）を裏付けるものであり、林大氏旧蔵の日下部表こそが選定作業の台帳であることがわかる。

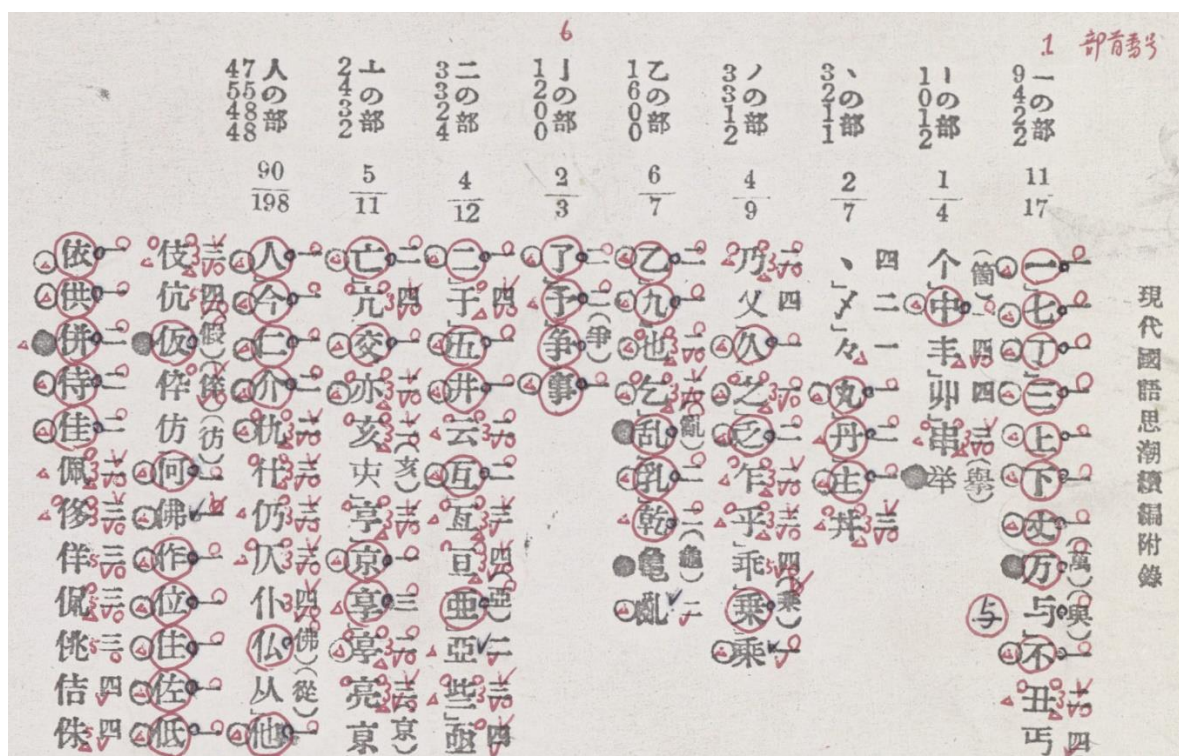


図 7: 日下部表（一部）

付. 国立国語研究資料室

国立国語研究所は、1948年設立以降に実施した調査研究の資料を整理・保存する組織・設備として、共同利用推進センター内に研究資料室を設置している。調査票、録音・録画、録音・録画の文字化文書、情報カード、集計表、調査計画書、会議録など、主に研究成果である報告書や論文、データベース等にまとまる以前の、調査研究遂行段階での収集物や中間生成物を集積し、収蔵資料の各種目録を作成するとともに（国立国語研究所研究資料室収蔵資料 <https://rnr.ninjal.ac.jp/>）、調査・研究・教育を目的とする外来利用者や、国立国語研究所の共同研究プロジェクト（公募研究を含む）に収蔵資料を提供している。本発表で取り上げた資料は、研究資料室の収蔵資料である。

参考文献

- 木村拓也（2006）「戦後日本において『テストの専門家』とは一体誰であったのか？ —戦後日本における学力調査一覧と『大規模学力テスト』の関係者一覧—」『教育情報学研究』4、pp.67-100
- 佐藤寧（2014）「終戦直後に実施された世論調査の再検証」『日本世論調査協会報「よろん」』114、pp.26-37
- 読み書き能力調査委員会（1951）『日本人の読み書き能力』、東京大学出版部
- 国立国語研究所（1968）『国立国語研究所報告 31 電子計算機による国語研究』、秀英出版
- 甲斐睦郎（2004）「追悼林大先生」『国語学』55-4、pp.3-6
- 芝野耕司（2002）「漢字・日本語処理技術の発展：漢字コードの標準化」『IPSJ Magazin』43-12、pp.1362-1367
- 池田証寿（2004）「JIS 漢字第 1 次規格開発に対する林大氏の貢献」『国語学』55-4、pp.7-10
- 高田智和（2012）「日下部重太郎の漢字研究—三種の「日下部表」と『國民字典』—」『新時代的世界日語教育研究』（北京・高等教育出版社）、pp.231-239