

[学術論文]

日本の苗字の漢字の特徴

——ウェブサイトのデータからわかること——

Characteristic Traits of Kanji in Japanese Surnames as Based on Online Data

成田 徹男
NARITA, Tetuo

0. はじめに
1. ウェブサイトについて
 - 1.1 インターネットサイト「同姓同名辞典（全国の苗字ランキング）」のデータの特徴
 - 1.2 「同姓同名辞典（全国の苗字ランキング）」「1～10,000位」データから何がわかるか
2. 集計結果とそこに見られる傾向について
 - 2.1 苗字の文字数について——二文字がおよそ9割
 - 2.2 苗字に使われる文字の、延べと異なり
 - 2.3 苗字に使われる文字のありかたについて
 - 2.3.1 「1～10,000位」に使われた漢字 ベスト20
 - 2.3.2 漢字ではない、とみなしうる文字について
 - 2.3.3 使用数20以上のもの（21位～224位、207字）
3. おわりに

要旨 本稿では、インターネットサイト「同姓同名辞典（全国の苗字ランキング）」の「1位～10,000位」で公開されている表記データをもとに、日本の苗字の文字表記、主として漢字表記の特徴を整理してみた。

約10,000種の日本の苗字について、文字表記として次のような特徴があることがわかった。

- (1) 二文字苗字が約9割であること
- (2) 「ツ」「ノ」という片仮名が使われる苗字があること
- (3) 異なりはさほど多くなく、1, 155であったこと
- (4) 文字によって、〈頭〉〈中〉〈尾〉という分布に偏りがあるものがあること
- (5) 「田」がもっとも多く使われていること
- (6) よく使われる漢字は「藤」を除き、画数があまり多くない傾向があること
- (7) 上位20で使用数全体の3割弱、上位2割弱で8割強占めていること
- (8) 漢数字には〈頭〉タイプが多いこと

キーワード：日本の苗字、漢字、異なり、〈頭〉タイプ、〈中〉タイプ、〈尾〉タイプ

0. はじめに

日本の苗字に関するインターネットサイトは数多くある。数千以上のランキングを示すものだけでも数種はあるが、いずれも、もととなるものは、いわゆる「電子電話帳」のデータである。「電子電話帳」で比較的新しいものとしては『電子電話帳 2012 Ver.17 全国版』（日本ソフト販売）があり、日本全国47都道府県における全電話帳データ約2636万件(個人法人ふくむ)が収録されているという。また、ソフトボックスの『写録宝夢巢 Ver.19』は、2016年度版最新電話帳データベース(2015年4月までの発刊)を収録しているそうである。ソフトボックスは「姓名分布&ランキング 写録宝夢巢」というウェブサイトで苗字のランキングを示しているが、2016年4月8日にアクセスした時点では「2007年10月までに発刊された全国の電話帳に掲載されている約2,338万世帯の情報をもとに、分布、ランキング情報を表示しています」ということであつた。

そのようなサイトのうち、ある程度まとまった数の苗字ランキングを明示し、表記について注意をはらっているものとして、今回は「同姓同名辞典(全国の苗字ランキング)」をとりあげた。全国ランキング上位10,000の表記データをもとに、日本の苗字の漢字表記について、どのようなことが言えるのか整理してみることにする。このサイトは、苗字だけでなく「下のなまえ」も対象としているという特徴をもっているが、今回は苗字のみを対象とする。

1. ウェブサイトについて

1.1 インターネットサイト「同姓同名辞典(全国の苗字ランキング)」のデータの特徴

インターネットサイト「同姓同名辞典(全国の苗字ランキング)」は、「苗字検索」の項目の情報では、2015年5月10日現在で、漢字で10万種掲載されているそうである。また、出典元について「同姓同名辞典で使用しているデータは、全国電子電話帳CD-ROM(2002年版)から抽出したもので、大元のデータはNTT発刊のハローページ(平成13年(2001年)発行)です。」と説明されていた。

1.2 「同姓同名辞典(全国の苗字ランキング)」 「1~10,000位」データから何がわかるか

この資料を扱うのに際して、前提として留意しておくべき点がいくつかある。

- (1) 電子電話帳の資料をもととしているため、本名であるかどうかはわからない。仮に「なりた☆てつお」という芸名かペンネームで電話帳に登録し、「なりた☆」を苗字とすれば、この表記が(最下位であるが)ランキングに示される可能性がある。
- (2) 電子電話帳の文字データに基づき、「よみ」には言及していない。「林」さんが、「はやし」さんなのか、「りん」さんなのか、判断のしようがない。
- (3) 電子電話帳の資料をもととしているため、その人の国籍は不明である。「林」さんが、

（日本国籍の可能性が高い）「はやし」さんなのか、日本国籍の「りん」さんなのか、台湾の「りん」さんなのか、判断のしようがない。また、例えば、薩摩焼の「沈壽官【ちん・じゅかん】」は、窯の名であり現在15代目の当主が襲名している陶芸家の名でもあるが、元は、豊臣秀吉の二度目の朝鮮出征の折に、薩摩の島津義弘によって連行された朝鮮人技術者の中の陶工「沈当吉」を祖とする家系であると言われている。現15代目の戸籍上の苗字は「大迫」であるらしい。したがって、もし「沈壽官」のような名前が電子電話帳にあったとしても、その人の国籍は不明である。

（4）順位は世帯数によるけれども、その世帯数は、あくまで推定である。

戸籍に掲載されているのと同じ文字表記のものだけとは限らないので、「日本人の苗字」とは言えないし、「日本の苗字」と言ってよいかどうかにも問題があるかもしれない。けれども、「日本に住んでいる人の名乗っている苗字」にはちがいない。電子電話帳に掲載される「日本に住んでいる人」の名前には、そもそも本名以外の通称もふくまれる可能性があり、その場合には、苗字といわゆる「下のなまえ」との区別ができない場合もあるかもしれない。さらに、漢字で表記していない、あるいは漢字で表記できない苗字もふくまれうる。戸籍法上「下の名前」では使えない、アルファベットなどが使われる可能性もある。ただし、漢字以外の表記は、後に述べるようにわずかな例外はあるものの、10,000位までにはほとんど出てきていない。

日本の苗字は30万あるともいわれるが、それは、意味・用法が同じでも形がちがう漢字を別とみなすかどうか、という、いわゆる異体字の問題と、同じ文字表記（要するに漢字表記）であっても複数ありうる「よみかた」の問題との二点について、すこしでもちがうものはすべて別のものとして数えれば、という前提にもとづいて考えられた数値である。筆者は苗字の種類について、実際には、およそ12万から多くても20万程度であろうと推定している。その数値が正しければ、10,000種類は、5%から8%程度に当たる。全体の傾向を見るに十分な量であると考えた。「珍しい」というのは主観的な判断であるが、10,000種類の苗字には、後で触れる「勅使河原」とか、「名嘉真」とか、「武笠」とか、「五郎丸」とか、たまたま知り合いでもない限りは「ありふれている」とはいいいにくい苗字もふくまれている。また、主たる関心は、苗字にはどのような文字が多く使われているのかを確認することにある。そのためには上位10,000種類の数値を確認することが重要である。使われている漢字のバリエーションの広がりや豊かさも興味深いのが、それについては、さらに下位までふくめた分析が必要になってくる。

参考までに、このサイトによる、苗字の推定世帯数ベスト20は、次のとおりである。

順位	苗字	推定世帯数
1	佐藤	446662
2	鈴木	394339
3	高橋	328106
4	田中	309340
5	伊藤	251059

6	山本	250495
7	渡辺	250010
8	中村	244921
9	小林	237059
10	加藤	198992
11	吉田	193911
12	山田	190224
13	佐々木	166452
14	山口	149437
15	松本	146787
16	井上	141042
17	木村	134699
18	斎藤	131689
19	林	124512
20	清水	122038

2. 集計結果とそこに見られる傾向について

2.1 苗字の文字数について——二文字がおおよそ9割

まず、「1~10,000位」の10,000種類の苗字の文字数である。一文字だけのものは329、二文字のものが、予想がつくように一番多く8911、三文字のもの756、四文字のもの4であった。単純に考えれば、二文字苗字が89%強を占め、一文字苗字は3%強、三文字苗字が7%強、四文字苗字は0.04%となる。二文字苗字が9割の圧倒的多数で、一文字と三文字で1割程度、四文字以上はきわめて少数派、ということである。もちろん、下位にいくほど、いわば少数派の苗字が多くなるわけなので、いくらか割合は変動するかもしれない。なお「1~10,000位」にふくまれていた四文字のものの表記を示すと、「4026位 勅使河原」「6040位 勅使川原」「7331位 小比類巻」「8657位 大豆生田」である。おそらく、前ふたつは「てしがわら」、三つ目は「こひるいまき」四つ目は「おおまみゅうだ」ないし「おおまめうだ」と「よむ」のがふつうであろう。

一文字だけの苗字の表記を上位から順に示すと、次のとおりである。「林、森、原、東、辻、関、堀、南、岡、西、谷、星、堤、泉、岸、菅、畑、橘、平、柳、牧、島、秦、乾、奥、角、北、今、中、浜、楠、沖、滝、湊、浦、境、坂、金、伴、沢、神、梶、榊、堺、藤、長、仲、芝、椿、城、桂、佃、脇、新、塙、柴、磯、巽、高、峯、嶋、宮、迫、榎、向、峰、柏、杉、武、所、前、近、池、安、荒、表、栄、館、清、盛、宗、丹、鶴、李、澤、郡、轟、渡、館、畠、里、室、俵、上、郷、笥、泊、勝、槇、幸、旭、丸、源、間、藪、鏡、門、窪、笹、昆、山、龍、井、笠、進、迎、篠、黛、碓、張、橋、峠、坪、下、倉、守、本、薄、濱、麓、内、萩、嵐、瀧、奥、叶、沼、荻、

英、越、開、登、朴、園、枝、蒲、陳、吳、戎、福、彙、川、扇、斎、田、広、潮、管、勇、県、村、楯、徳、裕、後、幅、重、梯、松、坪、町、港、釣、富、庄、都、栢、綾、大、米、久、竜、端、嶺、阪、竹、住、祝、熊、巴、捧、石、空、亘、楨、岩、蔵、静、豊、元、稻、王、鄭、台、納、隅、蔀、霞、恵、甲、植、洞、萬、洌、吉、寿、末、蔦、要、政、嶽、段、野、錦、車、灘、番、黄、団、紀、網、縣、万、塘、永、蘭、藪、財、鷺、昇、保、磯、鴨、続、願、明、塩、檀、坊、菊、杠、駒、青、秋、定、胡、崔、河、積、劉、時、淵、縄、壇、斧、光、伝、小、釜、茂、鈴、益、桜、行、楓、巖、淀、徐、落、京、塚、梁、庵、崎、隈、成、宝、正、椎、水、楊、埤、猪、道、頼、孫、墨、趙、翁、畝、莊、市、善、桐、江、達、寺、貞、雫、代、工、蓮、延、文、蕨、梅、鉄、菌、寛、玉、実、白、富、流、目、因、廣、細、綿

苗字のベスト20にも入る19位の「林」を筆頭に、おなじみのものもある。「畑」「峠」「坏」などいわゆる「国字」がかなり見られるほか、中国や朝鮮の苗字と思われるものもある。しかし、先にも述べたとおり、「よみ」は不明であり、国籍も不明である。

2.2 苗字に使われる文字の、延べと異なり

一文字329、二文字8911、三文字756、四文字4なので、延べの字数は「 $329 + 8911 \times 2 + 756 \times 3 + 4 \times 4$ 」で、20,435である。異なりは1,155であった。一文字当たりの使用数は、平均17.7回ということになる。使用数5以下は、使用数5が51字（4.4%）、使用数4が65字（5.6%）、使用数3が98字（8.5%）、使用数2が170字（14.7%）、使用数1が309字（26.8%、つまり4分の1以上）であった。使用数5以下の累計は693字（60.0%）であった。6割は使用数5以下ということである。

異なりは意外に少なく、千あまりの文字を組み合わせることで、かなりの苗字が表記できるのである。その一方で、2.1.で見たように、一文字苗字には、常用漢字表にないような漢字もかなりふくまれている。それを考え合わせると、幅広くよく使われる漢字と、一般的な日本語の文章にはあまり使われないが、苗字（それから、たぶん地名）のような固有名詞に限って使われる漢字とがあると考えられる。たとえば今回の資料で使用数5以上の漢字では「藪」「彙」「菱」「槻」「湊」だとか、使用数1では「醜」「翻」「纈」「纈」だとかがこの後者にあたる。これらの多くは「知らなくてもさほど困らない」が、「知らないと読めない」ものである。

2.3 苗字に使われる文字のありかたについて

2.3.1 「1～10,000位」に使われた漢字 ベスト20

「山田」「田中」には漢字「田」が使われている。この二つの苗字の範囲で、漢字の異なりは「田」「山」「中」の3つで、使用数は、「田」が2、「山」「中」が1である。このように見たとき、文字の異なりはどれぐらいあるのだろうか。「1～10,000位」の10,000種類の苗字に使われた、その文字の苗字に使用された数を数えてみた。いわば文字の使用頻度一覧を作成したのである。同一漢字がふたつ続けて繰り返される例があるか、「山田山」「田中田」のように同一

漢字が2回使われている苗字があるか、についてはおそらくと思われるが、確認はできなかった。「13位 佐々木」や「野々山」「多々良」のような「々」を使った苗字は、かなりあった。その一方で10,000位までには「佐佐木」はなかった。「々」については、後で「2.3.2 漢字ではない、とみなしうる文字について」において触れる。

まず、上位20の漢字を次ページの【表1】に示した。「田」が、断然第一位で、729の苗字で使われていた。総文字数20,435字の3.57%にあたる。以下、「野」「川」「山」「谷」と自然に関わる漢字が上位にあらわれている。予想どおりと言えそうだが、異体字の問題がからむと、苗字の順位はすぐに大幅に変動することになるので、そこには注意を払わなければならない。例えば、「野」と「埜」と「乃」と「之」という四つの漢字は別に数えるのか、それならば「岡」と「丘」とは別なのか、「沢」と「澤」はどうなのか、「島」と「嶋」と「鳶」はどうするのか、「富」と「冨」は区別するのか、など、さまざまな問題があることを踏まえなければならない。あくまで元のデータで同じ漢字とされ、ひとつの漢字表記であらわされているものを数えた、という制約が前提として存在する。このサイトでは「旧字体について」として、次のような説明がなされていた。

元データには苗字だけで使用される旧字体の「さき」や「たか」などがふくまれています。このような文字は一部のパソコンでは文字化けして見られませんが、新字体の「崎」や「高」に変換をしています。したがって、この旧字体がふくまれる苗字での検索やランキングはありませんので、ご了解ください。逆にこの旧字体を変換した新字体を使っている苗字のランキングはアップする形になります。

なお、パソコンで文字化けをしない旧字体の「澤」や「濱」は、「沢」や「浜」と区別してデータのまま使用しています。(2015年5月10日現在)

「崎」「高」は「旧字体」ではなく、現在の日本では、ほぼ苗字など固有名詞専用の特殊な字体と考えるべきであるが、それらは区別されていない、ということである。このサイトでは、JIS漢字として区別されているものは、基本的に別のものとみなす、という方針なのである。したがって、「沢」と「澤」(これは旧字体)はもちろん、上記で例示した、「野」と「埜」と「乃」と「之」、「岡」と「丘」、「島」と「嶋」と「鳶」、「富」と「冨」も、すべて区別している。ただし、「辻」については、常用漢字表改訂以降、新しいJIS漢字として「二点しんにゅう」の字体を採用しているので、戸籍上は「一点しんにゅう」であるという場合も、あるいはかつてはもっぱら「一点しんにゅう」で表記していたというような場合もあるにちがいないけれども、すべて「二点しんにゅう」に統一されているようである。

「野」「崎」が11画であるけれども、【表1】のほとんどの漢字は画数が10画以内である。その中で、「藤」だけが18画で、目立って複雑な漢字である。これは、むしろその多くが「藤原氏」に由来したり関係したりする、漢字「藤」を使った苗字がたくさんあるからである。「藤原」

は46位であるが、「藤」が前に使われているものでも100位以内に、「30位 藤田」「40位 藤井」「77位 藤本」がある。先に示したベスト20では「1位 佐藤」をはじめ「5位 伊藤」「10位 加藤」「18位 斎藤」があるし、100位以内に「32位 後藤」「37位 近藤」「39位 遠藤」「44位 斉藤」「67位 工藤」「72位 安藤」がある。

【表1】 「1～10,000位」に使われた文字 ベスト20

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
1	田	729	97	13.3%	40	5.5%	592	81.2%
2	野	445	74	16.6%	43	9.7%	328	73.7%
3	川	403	93	23.1%	8	2.0%	302	74.9%
4	山	399	102	25.6%	4	1.0%	293	73.4%
5	谷	349	47	13.5%	13	3.7%	289	82.8%
6	井	336	76	22.6%	14	4.2%	246	73.2%
7	木	294	71	24.1%	22	7.5%	201	68.4%
8	本	280	42	15.0%	6	2.1%	232	82.9%
9	原	267	23	8.6%	7	2.6%	237	88.8%
10	小	261	259	99.2%	2	0.8%	0	0.0%
11	大	245	245	100.0%	0	0.0%	0	0.0%
12	村	222	45	20.3%	1	0.5%	176	79.3%
13	島	216	36	16.7%	0	0.0%	180	83.3%
14	中	203	140	69.0%	2	1.0%	61	30.0%
15	岡	194	44	22.7%	0	0.0%	150	77.3%
16	沢	192	27	14.1%	0	0.0%	165	85.9%
17	崎	183	11	6.0%	0	0.0%	172	94.0%
18	上	182	100	54.9%	0	0.0%	82	45.1%
19	藤	176	72	40.9%	2	1.1%	102	58.0%
20	松	167	84	50.3%	3	1.8%	80	47.9%

なお、ベスト20の累計使用数は5,733で、全文字数20,435の28%強に当たる。上位20が使用数全体の3割弱を占めているわけである。

【表1】の右に示してある数値について説明する。苗字のどの位置に使われているかによって分類して集計し割合を示してみた。〈頭〉とは、「一文字苗字」もふくめて苗字の一文字目に使われていた数である。つまり、その文字が一文字目に使われていた「一文字苗字」の数+「二文字苗字」の数+「三文字苗字」の数+「四文字苗字」の数である。〈尾〉とは、「一文字苗字」

を除き苗字の最後に使用されていた数である。つまり、その文字が二文字目に使われていた「二文字苗字」の数+三文字目に使われていた「三文字苗字」の数+四文字目に使われていた「四文字苗字」の数である。〈中〉とは、上記の〈頭〉〈尾〉ではないもの、すなわち、その文字が二文字目に使われていた「三文字苗字」の数+二文字目に使われていた「四文字苗字」の数+三文字目に使われていた「四文字苗字」の数である。「一文字苗字」は、〈頭〉とも〈尾〉とも（あるいは〈中〉とも）みなすことはできるが、ここでは〈頭〉とみなした。

「田」は、〈尾〉が8割以上であるが、〈頭〉でも〈中〉でも使われる。「野」は、それに比べると〈尾〉の比率が下がり、〈頭〉と〈中〉が増えている。「四文字苗字」4例、「勅使河原」「勅使川原」「小比類巻」「大豆生田」には「野」は使われていないので、「野」の約1割、43例は、「三文字苗字」の中央で使われていることになる。

10位の「小」は〈尾〉が0%、〈中〉が0.8%で、〈頭〉が99.2%、11位の「大」は〈尾〉も〈中〉も0%で、〈頭〉が100%である。つまり、このふたつは、用法が〈頭〉に極端にかたよっている。文字によって〈頭〉、〈中〉、〈尾〉への偏りによる区別が、ある程度見られるのである。

2.3.2 漢字ではない、とみなしうる文字について

今回のデータの文字をJISコード順に並べると、初めの方に「ケ」14例、「ツ」7例、「ト」1例、「ノ」20例が出てくる。

「ケ」は「3281位 鐘ケ江」「3357位 池ケ谷」「4388位 矢ケ崎」「5347位 矢ケ部」「5737位 竹ケ原」「5983位 西ケ谷」「6144位 三ケ尻」「6941位 針ケ谷」「7737位 桐ケ谷」「8074位 戸ケ崎」「8283位 梶ケ谷」「8480位 藤ケ崎」「8614位 安ケ平」「9161位 三ケ田」に使われていた。おそらく「が」ないし「か」と発音されるものが大部分であろう。この字については諸説あるが、文法的に言えば連体格助詞と思われ、漢字「箇」に相当するものと考えておく。つまり、漢字としてあつかっておく。

「ツ」は「4350位 三ツ木」「4510位 三ツ井」「6220位 三ツ橋」「7759位 四ツ谷」「8325位 二ツ森」「9645位 三ツ石」「9012位 三ツ谷」で使われていた。表記上なくても、たとえば「三井」は「みつい」と読まれることがあるので、「みい」ではなく「みつい」だと読みを限定する注記の役割を果たしている、とみなすことができる。オトを注記する、片仮名としてあつかっておく。つまり、苗字には漢字以外の文字も使われているのである。

「ト」は「3346位 ト部」に使われている。おそらく一般的には「うらべ」と読み、「ト」については「占」と同じ意味をあらわす漢字としてあつかっておく。

「ノ」は「1344位 一ノ瀬」「3009位 井ノ口」「3901位 二ノ宮」「4241位 竹ノ内」「4765位 井ノ上」「5667位 山ノ内」「6021位 一ノ宮」「6269位 藤ノ木」「6881位 木ノ下」「6937位 山ノ井」「7003位 五ノ井」「7169位 堀ノ内」「7303位 木ノ内」「7319位 田ノ上」「7958位 四ノ宮」「8347位 池ノ谷」「8963位 池ノ上」「8095位 島ノ江」「9388

位 中ノ瀬 「9942位 一ノ関」で使われていた。連体格助詞の「ノ」であり、「ツ」と同じく、オトを注記する、片仮名としてあつかつておく。

また、J I Sコード順に並べた最後に登場するのが「々」である。漢字「仝」に由来するが、それ自体では対応するオトが定まらないので、文字研究者は「繰り返し符号」とし、文字とは区別する。しかし、漢字に後接して使われ、前の漢字と複合して複合語ないし複合語の一部をつくるので、使用者の感覚としては「漢字」のひとつである。ここでは漢字としてあつかつておく。29の苗字例を示す。「佐々」「百々」「野々」「佐々木」「野々村」「野々山」「多々良」「野々垣」「加々美」「佐々野」「野々下」「野々口」「千々岩」「佐々本」「等々力」「佐々田」「千々和」「佐々井」「多々納」「野々上」「野々宮」「間々田」「野々部」「加々見」「佐々岡」「須々木」「千々松」「野々川」「小佐々」。

要するに「ツ」と「ノ」は、例外的に苗字に使われる片仮名である。「水の江瀧子」が芸名を戸籍名にしたそうなので、今回の資料にはなかったが、平仮名「の」も使われているようだ。

2.3.3 使用数20以上のもの（21位～224位、207字）

【表1】に続く、21位から、使用数20の、同列224位の4つの文字まで、計207字を示すと、次のとおりである。「順位 文字 使用数」というかたちで示す。

「21位 森 148」	「22位 尾 147」	「23位 高 146」	「24位 石 143」	「24位 内 143」
「26位 久 142」	「27位 下 138」	「28位 江 130」	「29位 坂 126」	「29位 西 126」
「31位 永 124」	「31位 宮 124」	「33位 戸 120」	「34位 瀬 119」	「35位 吉 118」
「35位 倉 118」	「37位 津 114」	「38位 部 113」	「39位 平 112」	「40位 三 111」
「41位 安 109」	「42位 澤 108」	「43位 間 107」	「44位 見 100」	「45位 口 99」
「46位 古 98」	「47位 長 95」	「48位 岩 93」	「49位 佐 90」	「50位 竹 87」
「51位 橋 86」	「52位 河 81」	「52位 池 81」	「52位 嶋 81」	「55位 元 79」
「55位 水 79」	「57位 保 77」	「58位 根 75」	「59位 地 74」	「59位 塚 74」
「61位 金 73」	「62位 福 72」	「63位 屋 71」	「63位 北 71」	「65位 浦 70」
「65位 城 70」	「65位 畑 70」	「68位 新 69」	「68位 神 69」	「70位 武 68」
「71位 伊 67」	「71位 林 67」	「73位 矢 65」	「74位 沼 63」	「75位 生 59」
「76位 丸 58」	「76位 寺 58」	「76位 重 58」	「76位 前 58」	「76位 東 58」
「76位 浜 58」	「76位 堀 58」	「76位 名 58」	「84位 越 57」	「84位 賀 57」
「84位 出 57」	「87位 和 56」	「88位 富 55」	「89位 羽 53」	「89位 垣 53」
「91位 場 52」	「91位 日 52」	「93位 海 51」	「93位 広 51」	「93位 国 51」
「96位 波 49」	「97位 園 48」	「97位 今 48」	「97位 脇 48」	「100位 稲 47」
「100位 光 47」	「100位 成 47」	「100位 仲 47」	「100位 白 47」	「100位 里 47」
「106位 須 46」	「106位 柳 46」	「108位 杉 45」	「108位 末 45」	「110位 花 44」
「110位 黒 44」	「110位 赤 44」	「110位 八 44」	「110位 米 44」	「110位 有 44」

「116位 奥 43」 「116位 横 43」 「116位 関 43」 「116位 子 43」 「116位 道 43」
 「116位 徳 43」 「122位 加 41」 「122位 家 41」 「122位 真 41」 「122位 南 41」
 「126位 市 39」 「126位 渡 39」 「126位 馬 39」 「126位 友 39」 「126位 良 39」
 「131位 宇 38」 「131位 塩 38」 「131位 船 38」 「131位 代 38」 「135位 角 37」
 「135位 多 37」 「135位 門 37」 「135位 立 37」 「139位 清 36」 「139位 土 36」
 「141位 阿 35」 「141位 若 35」 「141位 方 35」 「144位 泉 34」 「144位 滝 34」
 「146位 一 33」 「146位 手 33」 「146位 秋 33」 「146位 勝 33」 「146位 浅 33」
 「146位 鳥 33」 「152位 熊 32」 「152位 住 32」 「154位 岸 31」 「154位 玉 31」
 「154位 相 31」 「154位 辻 31」 「154位 留 31」 「159位 栗 30」 「159位 富 30」
 「159位 渕 30」 「162位 近 29」 「162位 志 29」 「162位 室 29」 「162位 二 29」
 「162位 片 29」 「162位 々 29」 「168位 居 28」 「168位 月 28」 「168位 向 28」
 「168位 荒 28」 「168位 笹 28」 「168位 鹿 28」 「168位 深 28」 「168位 千 28」
 「168位 猪 28」 「168位 辺 28」 「168位 目 28」 「179位 溝 27」 「179位 阪 27」
 「179位 枝 27」 「179位 入 27」 「179位 飯 27」 「179位 牧 27」 「185位 笠 26」
 「185位 草 26」 「185位 増 26」 「185位 鶴 26」 「185位 天 26」 「190位 貝 25」
 「190位 細 25」 「190位 星 25」 「190位 町 25」 「190位 豊 25」 「195位 喜 24」
 「195位 窪 24」 「195位 合 24」 「195位 青 24」 「195位 梅 24」 「195位 迫 24」
 「195位 満 24」 「202位 磯 23」 「202位 兼 23」 「202位 五 23」 「202位 後 23」
 「202位 柴 23」 「202位 守 23」 「202位 之 23」 「202位 板 23」 「202位 比 23」
 「202位 利 23」 「212位 亀 22」 「212位 郷 22」 「212位 桑 22」 「212位 菅 22」
 「212位 明 22」 「217位 柿 21」 「217位 宗 21」 「217位 端 21」 「217位 堂 21」
 「217位 美 21」 「217位 淵 21」 「217位 峰 21」 「224位 ノ 20」 「224位 梶 20」
 「224位 形 20」 「224位 植 20」

なお、21位～224位の累計使用数は10,702で、全文字数20,435の52.4%に当たる。ベスト20の累計使用数5,733と合わせると1位～224位の累計使用数合計は16,435となり、これは全文字数の80.4%に当たる。上位224(1,155の19.4%)で使用数全体の80%以上をカバーしている。つまるところ、上位2割弱で8割強占めているわけである。

2.3.1 で述べた、〈頭〉、〈中〉、〈尾〉への偏りによる区別は、ある程度の使用数がないとわからない。ここでは使用数に占める割合が、90%以上である区分があるとき、その区分の名をつけて、「〈頭〉タイプ」、「〈中〉タイプ」、「〈尾〉タイプ」と呼ぶことにしておく。二文字苗字が9割なので、「〈中〉タイプ」と認定できるものはかなり限定されると予想されるが、ここではその点については議論しない。

21位～224位について、特徴的なものをみておこう。

【表2】 「21～224位」に使われた文字で特徴的なもの1 漢数字（〈頭〉タイプ）

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
40	三	111	110	99.1%	0	0.0%	1	0.9%
110	八	44	42	95.5%	2	4.5%	0	0.0%
146	一	33	32	97.0%	0	0.0%	1	3.0%
162	二	29	27	93.1%	2	6.9%	0	0.0%
168	千	28	28	100.0%	0	0.0%	0	0.0%
202	五	23	23	100.0%	0	0.0%	0	0.0%

「2本」「5杯」のような「数詞+助数詞」用法はもちろん、「竹林の七賢」「三美人」「五大臣」のような「数詞+名詞」用法でも数詞が先行する。「六本木」「三軒茶屋」などの地名に見られるように「数詞+助数詞+名詞」の順で複合語がつけられることもある。そんなところから漢数字をふくむ苗字で数詞が最初におかれる傾向があることは首肯できる。なお、「一～九」のうち、上記以外の漢数字は、「228位 四 19」「463位 六 5」「340位 七 10」「463位 九 5」であった。しかもこれらは、使用数は20未満ではあるものの、いずれも〈頭〉100%であった。

【表3】 「21～224位」に使われた文字で特徴的なもの2 〈中〉タイプ

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
202	之	23	0	0.0%	23	100.0%	0	0.0%
224	ノ	20	0	0.0%	20	100.0%	0	0.0%

<参考>

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
162	々	29	0	0.0%	25	86.2%	4	13.8%
280	ケ	14	0	0.0%	14	100.0%	0	0.0%
388	ツ	7	0	0.0%	7	100.0%	0	0.0%

「之」は、前述した「ノ」と似ていて、〈中〉タイプである。分布だけからすると、このふたつは（異なる文字種でありながら）異体字と見ることもできる。また、これも前に言及した「々」は、「佐々」など〈尾〉の例もあるため、〈中〉が90%に達しないが、〈中〉タイプとみなすことも可能である。もともと3文字以上の苗字でなければ〈中〉と認定できる文字用例がないわけなので、使用数20未満まで広げてみても、〈中〉タイプはこれも前に言及した「ケ」「ツ」ぐらいである。もう少し、個別に考察してみる必要があると思われる。

【表4】 「21～224位」に使われた文字で特徴的なもの3 漢数字以外の〈頭〉タイプ

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
68	新	69	69	100.0%	0	0.0%	0	0.0%
71	伊	67	65	97.0%	0	0.0%	2	3.0%
97	今	48	48	100.0%	0	0.0%	0	0.0%
100	白	47	46	97.9%	0	0.0%	1	2.1%
110	赤	44	44	100.0%	0	0.0%	0	0.0%
110	有	44	44	100.0%	0	0.0%	0	0.0%
116	奥	43	41	95.3%	0	0.0%	2	4.7%
116	横	43	43	100.0%	0	0.0%	0	0.0%
122	加	41	39	95.1%	0	0.0%	2	4.9%
131	宇	38	36	94.7%	1	2.6%	1	2.6%
141	阿	35	34	97.1%	1	2.9%	0	0.0%
146	秋	33	30	90.9%	0	0.0%	3	9.1%
146	勝	33	30	90.9%	0	0.0%	3	9.1%
146	浅	33	30	90.9%	0	0.0%	3	9.1%
162	片	29	27	93.1%	0	0.0%	2	6.9%
168	荒	28	28	100.0%	0	0.0%	0	0.0%
168	笹	28	27	96.4%	0	0.0%	1	3.6%
168	猪	28	28	100.0%	0	0.0%	0	0.0%
179	飯	27	26	96.3%	1	3.7%	0	0.0%
190	細	25	25	100.0%	0	0.0%	0	0.0%
190	豊	25	25	100.0%	0	0.0%	0	0.0%
195	青	24	24	100.0%	0	0.0%	0	0.0%
195	梅	24	23	95.8%	0	0.0%	1	4.2%

漢数字以外の〈頭〉タイプには、「白」「赤」「青」という色をはじめ形容詞的な(状態的な)意味をもつものがある。「青い山」のように名詞を修飾する形容詞は名詞に前接し、形容詞語幹と名詞で複合語を構成する場合も「青山」のように形容詞語幹が前接する。したがってこれらが前の方に位置するのは理由があることだと考えられる。それ以外については、今後、個別に考えてみたい。特に「伊」「加」は、「伊藤」「加藤」では地名をあらわしていると考えられるが、それ以外にも〈頭〉として多用される理由は何か、興味がある。

【表5】 「21～224位」に使われた文字で特徴的なもの3 〈尾〉タイプ

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
38	部	113	1	0.9%	1	0.9%	111	98.2%
42	澤	108	11	10.2%	0	0.0%	97	89.8%
45	口	99	1	1.0%	0	0.0%	98	99.0%
91	場	52	0	0.0%	2	3.8%	50	96.2%
141	方	35	1	2.9%	1	2.9%	33	94.3%
168	辺	28	2	7.1%	0	0.0%	26	92.9%
224	形	20	2	10.0%	0	0.0%	18	90.0%

＜参考＞（ベスト20の一部再掲）

順位	漢字	使用数	〈頭〉	〈頭〉%	〈中〉	〈中〉%	〈尾〉	〈尾〉%
16	沢	192	27	14.1%	0	0.0%	165	85.9%
17	崎	183	11	6.0%	0	0.0%	172	94.0%

「澤」は、90%未満なので厳密には〈尾〉タイプからは外れるのであるが、ここで言及しておきたい。16位の「沢」の旧字体なので、分布としては似ている。前接する漢字の一致度を調べれば「異体字」として認定する基準を示せるかもしれない。

3. おわりに

元のデータが電子電話帳であっても、サイトによって苗字のランキングはちがっていることがある。また、データの処理のしかた、整理のしかたによっては、この研究で示したような「漢字ランキング」でもちがった結論になる可能性はある。しかし、おそらく、「田」がもっともよく使われていることは間違いないし、「ノ」「ツ」という片仮名表記をふくむ多くの苗字が使われていることも間違いないであろう。

そういった点を、自分なりに自信をもって言えるようになっただけでも、ひとつの成果であったと思う。サイトには、「よみ」を示しているものがある。次のステップとして、苗字について、文字表記だけでなく、どうよむかというデータをもとに、何か新しい知見が得られないか模索しているところである。

参考文献・参考サイト

笹原宏之【ささはら・ひろゆき】2013『方言漢字』角川書店

笹原宏之【ささはら・ひろゆき】2014『漢字に託した「日本の心」』（NHK出版新書438）NHK出版

安岡孝一【やすおか・こういち】2011『新しい常用漢字と人名用漢字 漢字制限の歴史』三省堂

<http://www.douseidoumei.net> 「同姓同名辞典（全国の苗字ランキング）」 150520 アクセス

