

苗字・名前ランキング並びに地名データに於ける 音韻パターンの選択則

札幌学院大学 社会情報学部 早田和弥

要旨：統計的手法に拠り、標記日本語語彙データに於ける音韻パターンの解析を行った結果を示す。データ規模は、苗字と名前に対して共に三千、地名に対して約二千である。名前に就いては性別毎に解析し、地名に就いては全市町村並びに郡を対象としている。更に比較の為、日本酒銘柄及び枕詞に対しても同様な分析を行っている。単語中の母音配置を解析した結果、現代地名に就いてはこれが確率法則に従っていると判定されたが、苗字と名前及び古代郡名に就いては各々特徴的な分布を示した。これらのうち苗字と古代郡名に於いては、日本語基礎語彙の音韻パターンと基本的に一致することを確認したが、名前に於いては男女共これと著しく異なる極めて特異なパターンを生じた。即ち、男子名、女子名共に近接する母音間で同音回避が顕著であり、特に男子名に於いてこの傾向が強いことが分かった。

キーワード：語彙統計、氏名分析、地名学、音韻法則、枕詞

1 はじめに

苗字（名字）の起りは古代の氏姓制度にまで遡る。古代の姓は、職業、官職、地名、出身国などに由来するものが多いと言われているが、同一の氏から分れ出て、その住む地名などを取ったものがよく知られている。例えば、平氏から出た千葉・三浦の類がこれに当たる。江戸時代になって平民が苗字を唱えることを許され（苗字御免）、明治に入り現在の国民皆姓が布告され、現在に至っている。我が国に於ける苗字の総数は推定で15万とも27万とも言われており、正確な値は定かではない。今日、苗字を新たに開発し、本名としてこれを名乗ることは民法上許されていない。但し、現存する苗字の範囲内であれば、養子縁組や婚姻を通して変更が可能である。従って、「徳川」や「冷泉」などといった苗字を法律上名乗ることは（障壁は途轍もなく高いであろうが）絶対に不可能という訳ではない。一方、個人名を表す名前に就いては、苗字と異なり、新たなものの創生が原理上幾らでも可能である点に大きな特徴がある。然し乍ら、命名の際、苗字のように地名や出自を元にするとは一般的でない。所で、現在及び過去に存在した戸籍に記された名前は膨大な数に上るであろう。更に、記録に残っていないものをも含めた場合、即ち有史以前から現在まで我が国に存在した名前の累積は一体どれくらいに達するのであろうか。この問いは、命名の起源、更には言語の起源にまで遡る難問であり、その概算値を推定することさえも不可能であろう。さて、以上述べた苗字や名前といった人名とは対照的に、地名の総数は、小字以上の階層に限れば地図等を参照することに依って比較的容易に把握することができる。地名の変更・

消失は、市町村合併時或いは町村から市への移行時に於いて、今日でも屢々経験される⁽¹⁾。本論文の目的は、これらの名称データから音韻列を抽出し、これらが特定のパターンを選択しているか否か大規模統計解析を通して明らかにすることにある。以下では先ず、日本語基本語彙データを対象として音韻パターンの解析を行い、名詞を中心とした日本語一般の音韻構造を明らかにしている。引き続き、苗字、名前、地名といった名称データを解析し、日本語一般に対する結果との比較を行っている。データ規模は、苗字と名前に対して共に三千、地名に対して約二千である。ここに、名前に就いては男女別々に解析し、地名に就いては全市町村並びに郡を対象としている。更に比較の為、日本酒銘柄及び枕詞に対しても同様な分析を行っている。単語中の母音配置パターンを分類・解析することにより、これが確率法則に従っているか否かを判定することができる。

2 母音パターンの抽出と分類並びに解析

2.1 抽出法

単語の母音パターン抽出法を説明する為、苗字、男子名、女子名、地名の中から代表的なものとして、以下の四つを其々選ぶ。

たかはし しゅんすけ ゆうこ くねっぶ

これらの単語をローマ字で表記すると

Ta・ka・ha・shi Shun・su・ke Yû・ko Kun・nep・pu

となり、音節数は其々4, 3, 2, 3である(中黒は音節の区切れを表す)。今、各語から母音を取り出すと

aaai uue uo ueu

となる。ここに、各語の音節数と母音数は一致する。

2.2 分類法

「ゆうこ(uo)」のような2音節の単語では、{AB}と{AA}の2種類の音列に分けられる。ここにA, Bは異なる二つの母音を表す。従って、uoはABというパターンに分類されるが、別の例「みき(ii)」はAAに分類される。一方、「しゅんすけ(uue)」や「くねっぶ(ueu)」のような3音節の単語では、{ABC}, {AAB, ABA, ABB}, {AAA}からなる5種類の音列に分けられる。ここにCはA, Bとは異なる母音を表す。これに依ると、uueはAAB, ueuはABAに属する。他方、「たかはし(aaai)」のような4音節の単語の場合、以下に示す15種類のパターンが可能である。

{ABCD}, {AABC, ABAC, ABCA, ABBC, ABCB, ABCC}, {AAAB, AABA, ABAA, ABBB},
{AABB, ABAB, ABBA}, {AAAA}

ここにDはA, B, Cとは異なる母音を表す。これに依ると、aaaiは8番目のAAABというパターンに属するが、「きもべつ(ioeu)」町は1番目のABCDに、「やまがた(aaaa)」市は15番目のAAAAに分類される。

2.3 解析法

各音列パターンに属する実測度数の期待度数からの隔たりを χ^2 の値により評価する。ここに期待度数は組合せ論的確率の算出を通して求めることができる (Hayata, 2006)。実測度数の期待度数からの隔たりの非有意性 (偶発性) を帰無仮説 H とする。有意水準 α は、数表に従い 0.1%, 1%, 5% に定める。尚、 χ^2 分布の自由度は、規格化条件を考慮して 2 音節, 3 音節, 4 音節のデータに対して其々 1, 4, 14 となる (武藤, 1995: 418-419)。

3 結 果

3.1 日本語基礎語彙

先ず始めに、日本語の基礎語彙に対して本手法を適用し、その音韻パターンの特徴を明らかにする。語彙表として、安本による基礎500語 (安本, 1995: 185-198) に加え、比較の為、伊藤による同280語 (伊藤, 1994: 163-173) を考える。前者は全て我が国に於いて古来使用されている語から成り、名詞に加えて動詞、形容詞などの用言も可也の数に上っている。これに対し後者は、ポルトガル語、スペイン語などのロマンス語の基礎語彙を日本語に翻訳したもので、オリーブやワインなどといった日本語として必ずしも基礎的であると言えないものも含まれている。品詞と音節数の分布を其々表 1, 2 に示す。表 1 より、伊藤による語彙表では名詞が大部分を占め、動詞の割合が極めて低いことが分かる。又、表 2 より、音節数の分布に就いても両語彙表で大きく異なっている。音節数の最頻値は共に 2 で一致しているが、中央値は安本表では 2 であるのに対し、伊藤表では 3 となる。音節数が 2 と 3 の語彙に対する分類結果を其々表 3 と表 4 に示す。先ず表 3 (2 音節語) より

安本データに対して、 $\chi^2=0.14 < 3.841$ ($\alpha=5\%$)

表 1 : 日本語基礎語彙に於ける品詞の内訳.

品詞	安本データ		伊藤データ	
	度数	相対度数 (%)	度数	相対度数 (%)
名詞	297	59.4	241	86.1
動詞	160	32.0	19	6.8
形容詞	37	7.4	11	3.9
副詞	5	1.0	9	3.2
接続詞	1	0.2	0	0.0
合計	500	100.0	280	100.0

表 2 : 日本語基礎語彙に於ける音節数の分布.

音節数	安本データ		伊藤データ	
	度数	相対度数 (%)	度数	相対度数 (%)
1	26	5.2	19	6.8
2	274	54.8	108	38.6
3	163	32.6	98	35.0
4	35	7.0	44	15.7
5	2	0.4	9	3.2
6	0	0.0	2	0.7
合計	500	100.0	280	100.0

表 3 : 日本語基礎語彙 (2 音節) に於ける音韻パターンの比較. 母音分布は, 安本データに対して, a : 171, i : 109, u : 145, e : 41, o : 82, 伊藤データに対して, a : 52, i : 52, u : 52, e : 20, o : 40.

パターン	安本データ		伊藤データ	
	実測度数	期待度数	実測度数	期待度数
AB	207	209.63	84	84.59
AA	67	64.37	24	23.41
合計	274	274.00	108	108.00

表 4 : 日本語基礎語彙 (3 音節) に於ける音韻パターンの比較. 母音分布は, 安本データに対して, a : 147, i : 100, u : 157, e : 35, o : 50, 伊藤データに対して, a : 83, i : 66, u : 70, e : 18, o : 57.

パターン	安本データ		伊藤データ	
	実測度数	期待度数	実測度数	期待度数
ABC	62	63.24	45	41.68
AAB	46	29.44	28	16.98
ABA	30	29.44	14	16.98
ABB	10	29.44	7	16.98
AAA	15	11.45	4	5.37
合計	163	163.01	98	97.99

伊藤データに対して, $\chi^2=0.02 < 3.841$ ($\alpha = 5\%$)

となり, 共に帰無仮説 H は棄却できない. 故に, 両語彙表共に実測度数が確率法則に従っていることが分かる. 次に表 4 (3 音節語) より

安本データに対して, $\chi^2=23.29 > 18.466$ ($\alpha = 0.1\%$)

伊藤データに対して, $\chi^2=14.15 > 13.277$ ($\alpha = 1\%$)

が得られ, 其々有意水準 0.1%, 1% の下で本検定の有意性が明らかとなった. 表 4 に於いて実測度数と期待度数を比較すると, 双方共 AAB 型への偏りが顕著であることが分かる. 一方, これとは対照的に ABB 型への配分が際立って少なくなっている. 以上の解析結果から次の結論を下すことができる.

日本語の基礎語彙のうち 2 音節のものの音韻配列は確率法則に従っているが, 3 音節になると AAB 型への偏りが顕著になり, これと逆順の ABB 型への配分が少ない.

3.2 苗字ランキング

是迄公表されてきた, 苗字データに対する大規模な順位表の中で最も有名なものは佐久間ランキングであると言われているが, 刊行が昭和 39 年である為, 使用データが昭和 30 年代のものであると推定される. 以上の理由に拠り, ここでは最新の村山ランキング (村山, 2000) を採用する. 上位 1500 位並びに 3000 位迄の苗字デー

表 5 : 苗字ランキングに於ける音節数の分布.

音節数	1500位迄		3000位迄	
	度数	相対度数(%)	度数	相対度数(%)
1	7	0.5	15	0.5
2	189	12.6	379	12.6
3	663	44.1	1263	42.1
4	620	41.2	1284	42.8
5以上	26	1.7	59	2.0
合計	1505	100.1	3000	100.0

表 6 : 苗字ランキング (2 音節) に於ける音韻パターンの度数分布. 母音分布は, 1500位迄に対して, a : 131, i : 78, u : 27, e : 28, o : 114, 3000位迄に対して, a : 247, i : 154, u : 64, e : 66, o : 227.

パターン	1500位迄		3000位迄	
	実測度数	期待度数	実測度数	期待度数
AB	150	139.06	298	283.55
AA	39	49.94	81	95.45
合計	189	189.00	379	379.00

タに対する音節数及び音列分布を表 5 - 8 に示す. 先ず表 5 から, 音節数の中央値は共に 3 であり, その分布形状に就いても両者で殆ど違いが見られない. 敢えて言うならば, 後者に於いて僅かに音節数の大きい側へ偏っている. 2 音節データに対する分類結果を表 6 に示す. これより

上位1500位迄に対して, $\chi^2=3.26 < 3.841$ ($\alpha=5\%$)

上位3000位迄に対して, $\chi^2=2.92 < 3.841$ ($\alpha=5\%$)

を得る. 故に上記の日本語基礎語彙の場合と同様, 2 音節の苗字データに対しては本検定の有意性は主張できない. 次に 3 音節データに対する結果を表 7 に示す. これより

上位1500位迄に対して, $\chi^2=11.72 > 9.488$ ($\alpha=5\%$)

上位3000位迄に対して, $\chi^2=30.37 > 18.466$ ($\alpha=0.1\%$)

を得る. 因って 3 音節の苗字に対し本検定は有意水準が其々 5%, 0.1% の下で有意である. 表 7 に示されている実測度数と期待度数の比較から, 苗字データにおける音列パターンの配分は, 先に確認された日本語基礎語彙におけるものと本質的な違いがないことが分かる. 即ち, AAB 型への著しい偏りと BAA 型に対する回避傾向が苗字ランキングに於いても認められる. ここでは更に 4 音節データをも扱う. 音列分布を表 8 に示す. χ^2 の値は

上位1500位迄に対して, $\chi^2=122.28 \gg 36.124$ ($\alpha=0.1\%$)

上位3000位迄に対して, $\chi^2=216.24 \gg 36.124$ ($\alpha=0.1\%$)

表7：苗字ランキング（3音節）に於ける音韻パターンの度数分布。母音分布は、1500位迄に対して、a：856，i：456，u：231，e：106，o：340，3000位迄に対して、a：1565，i：912，u：434，e：220，o：658。

パターン	1500位迄		3000位迄	
	実測度数	期待度数	実測度数	期待度数
ABC	236	230.04	463	451.06
AAB	152	122.56	297	232.19
ABA	103	122.56	195	232.19
ABB	111	122.56	195	232.19
AAA	61	65.29	113	115.37
合計	663	663.01	1263	1263.00

表8：苗字ランキング（4音節）に於ける音韻パターンの度数分布。母音分布は、1500位迄に対して、a：1239，i：562，u：287，e：98，o：294，3000位迄に対して、a：2453，i：1199，u：630，e：207，o：647。

パターン	1500位迄		3000位迄	
	実測度数	期待度数	実測度数	期待度数
1) ABCD	30	44.57	62	101.17
2) AABC	55	44.94	117	97.12
3) ABAC	27	44.94	72	97.12
4) ABCA	38	44.94	87	97.12
5) ABBC	24	44.94	53	97.12
6) ABCB	47	44.94	90	97.12
7) ABCC	91	44.94	200	97.12
8) AAAB	26	46.06	73	90.00
9) AABA	53	46.06	100	90.00
10) ABAA	61	46.06	116	90.00
11) ABBB	40	46.06	79	90.00
12) AABB	52	27.01	94	56.28
13) ABAB	15	27.01	31	56.28
14) ABBA	15	27.01	33	56.28
15) AAAA	46	40.50	77	71.24
合計	620	619.98	1284	1283.97

となり、本検定の有意性が認められる。表8より次のことが分かる

- (1) 音列7) ABCCと音列12) AABBへの偏りが際立っている。
- (2) 音列1) ABCDと音列5) ABBCに対する回避傾向が強い。

3.3 名前ランキング

今世紀に入って、従来の苗字ランキングに加え、名前(男女名)ランキングが公表されるようになっていく。ここで採用するランキング表は、1998年から2000年迄の延3年間に、某月刊婦人雑誌に掲載された命名相談に関する記事などを元に推定されたものである(西沢・牧野, 2001)。男女共1000位迄と謳っているが、下位の方

で同順位の割合が増えてくるので、名前の収録数は其々1487, 1285に達し、共に1000を大幅に上回っている。最新の調査に基づいているので、古風な響きの名前は下位の方へ押し遣られ、現代風な名前や異国風な名前が取り分け女子名に於いて目を引いている⁽²⁾。男子名、女子名に対して、音節数並びに音列の度数分布を表9-12に示す。表の中には人名以外の欄（日本酒銘柄、枕詞）も併記されているが、これらに就いては後の第4節で触れたい。表9で最初に目を引くのは性別に依る音節数分布の著しい違いである。即ち、最頻値と中央値は男女共に3と成るが、その内容は全く異なっている。男子名に就いては、表5に示した苗字の音節数分布と類似した概形をもっている。これとは対照的に、女子名に於いては全体の99%が2音節と3音節に集中している。また5音節以上の長さをもつ女子名は皆無であり、1音節と4音節のものも1000位以内で其々僅かに八つ、五つしかない。以下にこれらを列挙する。先ず女子名で1音節のものは

ゆう (33), らん (259), あん (337), みい (491), りい (571), のん (677), みゅう (677), るう (677)

4音節のものは

さくらこ (270), ひまわり (297), かおるこ (677), こでまり (839), すみれこ (839)

となっている。ここに括弧内の数字はランキング中の順位を表す。これらの中で100位以内に入っているものは33位の「ゆう」のみである。名前ランキングに対する音列パターンへの配分結果を表10-12に示す。これらのうち、表10は2音節データに対する結果であり、 χ^2 値は

男子名に対して、 $\chi^2=13.24 > 10.827$ ($\alpha=0.1\%$)

表9：名前ランキング並びに日本酒銘柄に於ける音節数の分布。

音節数	男子名		女子名		日本酒銘柄	
	度数	相対度数(%)	度数	相対度数(%)	度数	相対度数(%)
1	25	1.7	8	0.6	0	0.0
2	288	19.4	376	29.3	125	6.8
3	639	43.0	896	69.7	201	11.0
4	518	34.8	5	0.4	390	21.3
5以上	17	1.1	0	0.0	1118	61.0
合計	1487	100.0	1285	100.0	1834	100.1

表10：名前ランキング（2音節）並びに日本酒銘柄（同）に於ける音韻パターンの度数分布。母音分布は、男子名に対して、a：115, i：120, u：122, e：49, o：170, 女子名に対して、a：234, i：188, u：115, e：73, o：142, 日本酒銘柄に対して、a：38, i：41, u：39, e：30, o：102。

パターン	男子名		女子名		日本酒銘柄	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
AB	250	223.93	303	290.35	93	93.10
AA	38	64.07	73	85.65	32	31.90
合計	288	288.00	376	376.00	125	125.00

表11：名前ランキング（3音節）並びに日本酒銘柄（同）に於ける音韻パターンの度数分布。母音分布は、男子名に対して、a：484，i：540，u：328，e：179，o：386，女子名に対して、a：843，i：710，u：385，e：216，o：534，日本酒銘柄に対して、a：124，i：162，u：119，e：80，o：118。

パターン	男子名		女子名		日本酒銘柄	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
ABC	336	281.14	422	374.04	104	93.12
AAB	89	108.12	184	155.89	19	32.90
ABA	110	108.12	163	155.89	39	32.90
ABB	82	108.12	84	155.89	28	32.90
AAA	22	33.50	43	54.27	11	9.17
合計	639	639.00	896	895.98	201	200.99

女子名に対して、 $\chi^2=2.42 < 3.841$ ($\alpha=5\%$)

となつて、性別に依つて異なる結果を齎す。即ち、男子名ランキングに対し本検定は有意水準0.1%の下で有意であること結論される。このような2音節語に対する検定の有意性は先の二つの場合には見出されなかった事実であり、注目に値する。これに対し女子名ランキングでは、既に見た日本語基礎語彙並びに苗字ランキングのときと同様、2音節語に対しては本検定の有意性は主張できない。次に3音節データに対する結果を表11に示す。これより

男子名に対して、 $\chi^2=24.38 > 18.466$ ($\alpha=0.1\%$)

女子名に対して、 $\chi^2=47.04 > 18.466$ ($\alpha=0.1\%$)

を得、2音節語のときとは異なり、判定結果は性別に依存しない。即ち、3音節語では両性に対し本検定は有意であることが示された(有意水準0.1%)。表11に於いて各音列パターンへの配分状態を比べてみると、AAB型以外は男女間で質的な差異は認められず同じ傾向が見られる。即ち、ABC型への強い偏向とABB, AAA型への回避が確認できる。就中女子名に於けるABB型への忌避は際立っている。然し乍ら、AAB型への配分は性別に強く依存している。具体的には、男子名に於いては本パターンへの集中が見られないが、女子名に於いては先の日本語基礎語彙と苗字の場合と同様、このパターンへの顕著な偏りが見られる。以上を要するに、男子名に対する結果は日本語語彙のものとは本質的に異なっていると言えるが、女子名に対する結果はABC型を除いてこれと定性的に同じ傾向にあると結論される。引き続き、4音節の男子名データに対する分類結果を表12に示す。これより

$\chi^2=46.07 > 36.124$ ($\alpha=0.1\%$)

となつて、有意水準0.1%の下で検定の有意性が立証された。表12の中の男子名に対する結果より次の事実を指摘できる。

- (1) 音列1) ABCDへの偏りが顕著である。
- (2) 同じグループ内の音列2)－7)を比較すると、2) AABCへの集中が見られる。
- (3) 同じグループ内の音列8)－11)を比較すると、8) AAABへの配分が最大で、これを反転したパター

表12：男子名ランキング（4音節）、日本酒銘柄（同）、及び枕詞（全5音節のうち、語尾の助詞1音節を除く）に於ける音韻パターンの度数分布。母音分布は、男子名に対して、a：539，i：589，u：334，e：166，o：444，日本酒銘柄に対して、a：429，i：408，u：317，e：125，o：281，枕詞に対して、a：446，i：286，u：246，e：89，o：229。

パターン	男子名		日本酒銘柄		枕詞	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
1) ABCD	106	78.14	58	60.05	42	44.79
2) AABC	66	48.17	43	36.44	56	29.47
3) ABAC	37	48.17	25	36.44	36	29.47
4) ABCA	50	48.17	32	36.44	21	29.47
5) ABBC	48	48.17	14	36.44	18	29.47
6) ABCB	45	48.17	33	36.44	12	29.47
7) ABCC	39	48.17	44	36.44	18	29.47
8) AAAB	31	21.33	18	15.69	25	14.74
9) AABA	15	21.33	17	15.69	12	14.74
10) ABAA	12	21.33	25	15.69	11	14.74
11) ABBB	11	21.33	18	15.69	20	14.74
12) AABB	14	19.43	29	14.45	20	12.47
13) ABAB	12	19.43	7	14.45	12	12.47
14) ABBA	28	19.43	18	14.45	13	12.47
15) AAAA	4	7.22	9	5.15	8	6.06
合計	518	517.99	390	389.95	324	324.04

ン11) ABBB への配分が最小となっている。

(4) 音列14) ABBA への偏りが見られる。

3.4 市町村名

周知のように、地名は世界のどの国でも階層構造を成しており、我が国では伝統的に、旧国→県→郡→市町村→字→小字と定められている。これらのうち冒頭の旧国は現在では採用されていない。本論文では市町村名を中心に郡名まで解析の対象とする。市町村名（矢野記念会：2007）に対する音節数の分布を表13に示す。市名と町村名の比較から、両分布は略一致していることが分かる。又、表5に示した苗字の分布と比較すると、表13では2音節語の相対度数が低くなっているのに対し、5音節語の相対度数が高くなっている。この特徴は、市町村名では「東・西・南・北」や「上・中・下」などといった方角に関する接頭辞、更には新十津川、新篠津のような「新」を冠する地名が頻出することに起因する。表13を見ると、市名・町村名共に1音節地名が唯

表13：地名データに於ける音節数の分布。

音節数	市名		町村名		市町村名	
	度数	相対度数(%)	度数	相対度数(%)	度数	相対度数(%)
1	1	0.1	1	0.1	2	0.1
2	171	21.9	178	17.4	349	19.3
3	283	36.2	410	40.1	693	38.4
4	229	29.3	304	29.7	533	29.5
5以上	98	12.5	129	12.6	227	12.6
合計	782	100.0	1022	99.9	1804	99.9

一つ存在するが、これらは其々、津市（三重県）と丹生町（福井県）に該当する。市町村名のうち2音節語の音列分布を表14に示す。これより

- 市名に対して、 $\chi^2=0.03 < 3.841$ ($\alpha = 5\%$)
- 町村名に対して、 $\chi^2=0.09 < 3.841$ ($\alpha = 5\%$)
- 市町村名に対して、 $\chi^2=0.00 < 3.841$ ($\alpha = 5\%$)

となり、どの場合も帰無仮説 H は棄却できない。即ち、日本語基礎語彙並びに苗字のときと同様に、2音節の市町村名の音韻選択は確率的であると結論される。次に、3音節語に対する配分結果を表15に示す。この表から

- 市名に対して、 $\chi^2=7.92 < 9.488$ ($\alpha = 5\%$)
- 町村名に対して、 $\chi^2=5.33 < 9.488$ ($\alpha = 5\%$)
- 市町村名に対して、 $\chi^2=4.91 < 9.488$ ($\alpha = 5\%$)

となり、2音節語のときと同様、帰無仮説 H は棄却できない。3音節語に対する本検定結果は、先に示した日本語基礎語彙、苗字、名前の場合とは異なるものであり、注目に値する。引き続き、4音節語に対する配分結果を表16に示す。 χ^2 値として

表14：地名データ（2音節）に於ける音韻パターンの度数分布。母音分布は、市名に対して、a：99，i：61，u：59，e：24，o：99，町村名に対して、a：112，i：74，u：48，e：20，o：102，市町村名に対して、a：211，i：135，u：107，e：44，o：201。

パターン	市名		町村名		市町村名	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
AB	130	130.97	136	134.28	266	265.52
AA	41	40.03	42	43.72	83	83.48
合計	171	171.00	178	178.00	349	349.00

表15：地名データ（3音節）に於ける音韻パターンの度数分布。母音分布は、市名に対して、a：332，i：190，u：120，e：51，o：156，町村名に対して、a：429，i：320，u：138，e：98，o：245，市町村名に対して、a：761，i：510，u：258，e：149，o：401。

パターン	市名		町村名		市町村名	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
ABC	109	107.38	173	162.31	282	270.83
AAB	57	50.97	81	73.02	138	123.80
ABA	62	50.97	57	73.02	119	123.80
ABB	39	50.97	73	73.02	112	123.80
AAA	16	22.71	26	28.64	42	50.77
合計	283	283.00	410	410.01	693	693.00

表16：地名データ（4音節）に於ける音韻パターンの度数分布。母音分布は、市名に対して、a：436，i：213，u：110，e：37，o：120，町村名に対して、a：552，i：282，u：140，e：64，o：178，市町村名に対して、a：988，i：495，u：250，e：101，o：298。

パターン	市名		町村名		市町村名	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
1) ABCD	18	18.21	28	27.95	46	46.23
2) AABC	22	17.39	41	24.01	63	41.52
3) ABAC	17	17.39	19	24.01	36	41.52
4) ABCA	10	17.39	16	24.01	26	41.52
5) ABBC	13	17.39	13	24.01	26	41.52
6) ABCB	13	17.39	21	24.01	34	41.52
7) ABCC	26	17.39	27	24.01	53	41.52
8) AAAB	14	15.96	18	19.71	32	35.65
9) AABA	12	15.96	16	19.71	28	35.65
10) ABAA	22	15.96	23	19.71	45	35.65
11) ABBB	21	15.96	20	19.71	41	35.65
12) AABB	14	10.02	26	12.89	40	22.88
13) ABAB	3	10.02	5	12.89	8	22.88
14) ABBA	14	10.02	11	12.89	25	22.88
15) AAAA	10	12.54	20	13.98	30	26.44
合計	229	228.99	304	303.50	533	533.03

市名に対して、 $\chi^2=24.55>23.685$ ($\alpha=5\%$)

町村名に対して、 $\chi^2=43.97>36.124$ ($\alpha=0.1\%$)

市町村名に対して、 $\chi^2=56.42>36.124$ ($\alpha=0.1\%$)

を得る。これより、市名に対しては有意水準5%で、町村名と市町村名に対しては0.1%で本検定の有意性が認められる。表16を元に4音節から成る市町村名の特徴を以下のように纏めることができる。

- (1) 音列2) AABCと12) AABBへの偏りが強い。
- (2) 音列4) ABCA, 5) ABBC並びに13) ABABに対する回避が見られる。

4 考 察

4.1 ロマンズ語基礎語彙

表3, 4に示した日本語基礎語彙に対する分析結果と比較する為、先に扱った伊藤データの基となっている5種類のロマンス諸語基礎語彙(伊藤, 1994: 163-173)に対して本手法を適用する。まず、単語の中に全部で2個の母音字を含む語彙に対する配分結果を表17に示す。本表より以下の値を得る。

ポルトガル語に対して、 $\chi^2=0.03<3.841$ ($\alpha=5\%$)

スペイン語に対して、 $\chi^2=0.01<3.841$ ($\alpha=5\%$)

フランス語に対して、 $\chi^2=4.26>3.841$ ($\alpha=5\%$)

イタリア語に対して、 $\chi^2=0.00<3.841$ ($\alpha=5\%$)

ルーマニア語に対して、 $\chi^2=3.22<3.841$ ($\alpha=5\%$)

表17：主要ロマンス諸語基礎語彙(母音字を二つ含む語)に於ける母音字パターンの比較。母音字分布は、ポルトガル語に対して、a：84, i：20, u：18, e：47, o：85, スペイン語に対して、a：88, i：22, u：18, e：46, o：68, フランス語に対して、a：31, i：48, u：25, e：92, o：44, イタリア語に対して、a：69, i：12, u：12, e：55, o：72, ルーマニア語に対して、a：47, i：26, u：36, e：28, o：19, ă：37, â+î：11.

(a) ポルトガル語, スペイン語, フランス語

パターン	ポルトガル語		スペイン語		フランス語	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
AB	94	93.11	89	89.40	100	90.23
AA	33	33.89	32	31.60	20	29.77
合計	127	127.00	121	121.00	120	120.00

(b) イタリア語, ルーマニア語

パターン	イタリア語		ルーマニア語	
	実測度数	期待度数	実測度数	期待度数
AB	80	79.87	92	85.29
AA	30	30.13	10	16.71
合計	110	110.00	102	102.00

これより、フランス語基礎語彙に於いてのみ検定の有意性が認められる(有意水準5%)。因ってフランス語の単語では、同じ母音字は近接して配置されない傾向にあることが分かる。これに対し、他の四つの言語に於いては、表3に示した日本語(2音節)のものと同じ判定結果を齎している。次に、単語中に全部で3個の母音字を含む語彙に対する結果を表18に示す。本表より以下の値を得る。

ポルトガル語に対して、 $\chi^2=22.46 > 18.466$ ($\alpha=0.1\%$)

スペイン語に対して、 $\chi^2=3.88 < 9.488$ ($\alpha=5\%$)

フランス語に対して、 $\chi^2=21.19 > 18.466$ ($\alpha=0.1\%$)

イタリア語に対して、 $\chi^2=4.99 < 9.488$ ($\alpha=5\%$)

ルーマニア語に対して、 $\chi^2=1.94 < 9.488$ ($\alpha=5\%$)

これより判定は言語により区別であることが分かる。先ず、検定の有意性(0.1%)が保証されるものは、ポルトガル語とフランス語であり、 χ^2 の値も両者間で大きな違いはない。更に、実測度数の分布状態も双方で似た傾向を示しており、共にABC型への集中が見られる。但し、両言語に於いても、日本語で見られたようなAAB型への集中やABB型に対する回避は存在しない⁽³⁾。

4.2 日本酒銘柄

前節で示した表9-12の中には、男子名と女子名に加え、標記銘柄(稲, 2007)の解析結果も併記されている。これらの表の中で最初に目に留まるのは、表9の音節数分布である。即ち、日本酒銘柄に於いては人名の場合とは異なり、音節数の増加に伴って相対度数が高くなり、音節数5以上のものが全体の6割以上を占めていることが分かる⁽⁴⁾。従って、音節数の中央値は5以上となる。この事実は、個人と商品に対する命名法の違いを考察する上で大変興味深い素材を提供している。人名の場合、随伴する苗字との整合性や背景となる伝統や文

表18：主要ロマンス諸語基礎語彙(母音字を三つ含む語)に於ける母音字パターンの比較。母音字分布は、ポルトガル語に対して、a：91, i：37, u：20, e：63, o：77, スペイン語に対して、a：114, i：40, u：34, e：87, o：79, フランス語に対して、a：47, i：56, u：37, e：104, o：41, イタリア語に対して、a：104, i：68, u：21, e：77, o：102, ルーマニア語に対して、a：52, i：54, u：26, e：52, o：19, ă：57, â+î：10。

(a) ポルトガル語, スペイン語, フランス語

パターン	ポルトガル語		スペイン語		フランス語	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
ABC	48	38.94	59	49.05	61	40.10
AAB	14	16.98	16	20.56	9	16.22
ABA	28	16.98	20	20.56	13	16.22
ABB	4	16.98	17	20.56	6	16.22
AAA	2	6.10	6	7.28	6	6.25
合計	96	95.98	118	118.01	95	95.01

(b) イタリア語, ルーマニア語

パターン	イタリア語		ルーマニア語	
	実測度数	期待度数	実測度数	期待度数
ABC	53	51.69	53	48.86
AAB	17	21.72	11	12.72
ABA	24	21.72	13	12.72
ABB	27	21.72	12	12.72
AAA	3	7.15	1	2.97
合計	124	124.00	90	89.99

化・習俗などから暗黙の制約を受け、無闇に語長を増やすことは許されない。これに対して、商標の場合には、先ず何よりも利益獲得などの経済的要因が優先される為、広告効果への期待感から名称が説明的になる嫌いがある。又、他社との競合も本傾向に拍車を掛けていると思われる。品格と伝統を重んじる日本酒銘柄に於いてもこのことは例外でなく、結果として土地柄や蔵元の特徴を色濃く反映した、より情報量の大きい名称が好まれることになる。表10に掲げた2音節語に対する分類結果から、次の値が導かれる。

$$\chi^2=0.00 < 3.841 \quad (\alpha = 5\%)$$

因って、女子名(2音節)のときと同様、日本酒銘柄(2音節)に於いても帰無仮説 H は棄却できない。尚、表11にある3音節語の分類結果を解析すると

$$\chi^2=9.37 < 9.488 \quad (\alpha = 5\%)$$

となつて、矢張り H は棄却できない。本判定結果は、前節で確認した男子名・女子名双方共異なっていることに注意されたい。一方、表12に見られる4音節の銘柄に対しては

$$\chi^2=64.93 > 36.124 \quad (\alpha = 0.1\%)$$

となつて、 H は棄却される。即ち、日本酒銘柄（4音節）の音韻パターンは確率法則に従っていないと判定される。パターン配分に於ける特徴は以下の通り。

- (1) 音列10) ABAA と12) AABB への偏向が認められる。
- (2) 音列5) ABBC と13) ABAB に対する回避が見られる。

4.3 枕詞

是迄見てきた4音節データの音韻パターン、具体的には苗字（表8）・男子名（表12）・市町村名（表16）・日本酒銘柄（表12）のパターン及び解析結果と比較する為、ここでは枕詞を考える。枕詞は、序詞や掛詞などと共に我が国上代文学に於ける修辞法として盛んに用いられたものであり、呪術から発したものだとか或いは語調を整える為だとか言われているが、その詳細な起源はよく解っていない。枕詞の総数は七百近くあるようであり、語の長さは殆どが5音節である⁽⁵⁾。修辞法としての第一の特徴は、特定の詞が直後の語（被枕詞）に掛かるといふ点にある。例えば、本居宣長が随筆集『玉勝間』巻八（本居、1934：317-318）で引用している二つの枕詞のうち、「ひさかたの（久方の、久堅の）」は天に関係した「天」「空」「雨」「月」「日」「昼」「雲」「雪」「霞」「月夜」などに掛かり、「あしひきの（足引の、葦引の）」は「山」「峰」「八峰」「尾の上」などに掛かる。本論文では、文献（内藤、2008）にある枕詞全678語のうち、5音節であつて且つ「名詞+助詞（の、を）」及び「動詞の終止形+助詞（や）」と分解できるもの、計324語を扱うことにする。上で記した二つは共に第一のタイプに属する。第二のタイプとして、例えば「あめしるや（天知るや）」がある。解析対象は、枕詞（5音節）のうち、語尾の助詞（1音節）を除いた4音節語とする。音列パターン（15種類）への分類を表12に示す。 χ^2 の値は

$$\chi^2 = 65.24 > 36.124 (\alpha = 0.1\%)$$

となり、帰無仮説 H は棄却される。表12から、枕詞の音韻パターンに関して次の事実が確認できる⁽⁶⁾。

- (1) 音列2) AABC, 8) AAAB 並びに12) AABB への偏りが認められる。就中音列2) 及び8) への集中が際立っている。即ち、枕詞全体では、語頭部に同音集中が見られる。この特徴は、前節で明らかとなつた日本語基礎語彙並びに苗字のそれを彷彿とさせるもので、強調に値する。
- (2) 音列5) ABBC, 6) ABCB, 7) ABCC に対する回避が見られる。この傾向は、特に音列6) に於いて著しい。

上記(1)で述べた性質をより詳しく考察する為、音列2) AABC と8) AAAB に焦点を絞り、 $A = \{a, i, u, e, o\}$ に対して両パターンの内訳を求めた結果を表19に示す。これより、両音列共、 $A = \{a, u, o\}$ に対して実測度数が期待度数より2乃至3倍高くなつてることが明らかとなつた⁽⁷⁾。尚、これら三種類の母音は全て母音三角形右側の同じ辺上（或いは母音空間の右側の領域）に位置しており興味深い。

4.4 郡名

前節では地名データとして市町村名を考えたが、本節では郡名に対する結果を示し、両者の違いを比較・考察する。周知の如く郡は町村に対する上位階層であり、行政単位としては大正時代に廃止されたが、今日でも地理上の区画として住所記載時などに慣用的に使われている。郡の歴史は非常に古く、大化の改新後に定められた国郡里制を発端とする。又、10世紀中盤に撰進され、本邦初の漢和辞書として知られる倭名類聚鈔（倭名

表19：枕詞(表12)に於ける音韻パターン2)AABC並びに8)AAABの内訳。BとCは、語頭部にある母音を除く互いに異なる母音を表す。

2) AABC			8) AAAB		
パターン	実測度数	期待度数	パターン	実測度数	期待度数
aaBC	29	11.88	aaaB	15	8.66
iiBC	4	6.58	iiiB	1	2.71
uuBC	12	5.29	uuuB	5	1.80
eeBC	0	0.97	eeeB	0	0.10
ooBC	11	4.75	oooB	4	1.47
合計	56	29.47		25	14.74

鈔，和名抄などと略)にも記載されていて，市町村名とは響きが異なる由緒ある名称が多いことでも知られる。例えば，鳳至(ふげし)郡(能登国→石川県)や肝属(きもつき)郡(大隅国→鹿児島県)などがこれに該当する。郡名簿(群馬大北川研：2009，日本地方行政研究会：1949，矢野記念会：1997，2009，郵政省：1999)に於ける音節数の推移を表20に示す。古代の郡名は和名抄記載の訓方に従っている。10世紀時点と現代では千年以上の隔りがあり，音節数の分布形状は大きく変化している。又，現代(1949-2010年)に於ける三つの分布を比べると，平成の大合併に伴う郡総数の大幅な減少にも拘らず，郡名長の分布は略一定で殆ど変化していないことが分かる。尚，現代に於いて見られる，音節数2並びに5以上をもつ郡の割合の顕著な増加の原因は，以下のように説明される。先ず前者に就いては，現代語に於ける音節数の短縮化傾向に由る。例として，甘楽：「かむら→かんら」，遠敷：「をにふ→おにゅう」，宍粟：「しさは→しそう」などが挙げられる。又，後者に就いては，主として方角を表す接頭辞の多用に由る。例えば，元々は越後国「蒲原郡」であったものが時代を経て新潟県「中蒲原郡」「東蒲原郡」「西蒲原郡」「南蒲原郡」「北蒲原郡」などと分化・拡大してゆくと，接頭辞の

表20：郡名データに於ける音節数分布の推移。

(a) 古代→昭和24年

音節数	和名抄記載訓		1949年	
	度数	相対度数(%)	度数	相対度数(%)
1	2	0.4	0	0.0
2	80	17.7	178	29.5
3	236	52.2	211	35.0
4	119	26.3	105	17.4
5以上	15	3.3	109	18.1
合計	452	99.9	603	100.0

(b) 平成9年→平成22年

音節数	1997年		2010年	
	度数	相対度数(%)	度数	相対度数(%)
1	1	0.2	1	0.3
2	163	29.5	104	27.0
3	188	34.0	133	34.5
4	106	19.2	78	20.3
5以上	95	17.2	69	17.9
合計	553	100.1	385	100.0

表21：郡名データ（2音節）に於ける音韻パターンの推移。母音分布は、和名抄記載訓に対して、a：55，i：34，u：27，e：15，o：29，1949年データに対して、a：127，i：78，u：54，e：21，o：76，1997年データに対して、a：123，i：68，u：49，e：21，o：65，2010年データに対して、a：82，i：41，u：31，e：14，o：40。

(a) 古代→昭和24年

パターン	和名抄記載訓		1949年	
	実測度数	期待度数	実測度数	期待度数
AB	59	61.33	128	133.97
AA	21	18.68	50	44.03
合計	80	80.01	178	178.00

(b) 平成9年→平成22年

パターン	1997年		2010年	
	実測度数	期待度数	実測度数	期待度数
AB	123	121.87	76	77.17
AA	40	41.13	28	26.83
合計	163	163.00	104	104.00

分だけ郡名の長さが増加することになる。長さが2音節である郡名に対する音列パターンの分布を表21に示す。これより

古代に於いて、 $\chi^2=0.38 < 3.841$ ($\alpha=5\%$)

1949年に於いて、 $\chi^2=1.08 < 3.841$ ($\alpha=5\%$)

1997年に於いて、 $\chi^2=0.04 < 3.841$ ($\alpha=5\%$)

2010年に於いて、 $\chi^2=0.07 < 3.841$ ($\alpha=5\%$)

となつて、年代に依らず帰無仮説 H は棄却できない。一方、長さが3音節である郡名に対する結果を表22に示す。解析結果は

古代に於いて、 $\chi^2=17.57 > 13.277$ ($\alpha=1\%$)

1949年に於いて、 $\chi^2=8.28 < 9.488$ ($\alpha=5\%$)

1997年に於いて、 $\chi^2=1.14 < 9.488$ ($\alpha=5\%$)

2010年に於いて、 $\chi^2=2.98 < 9.488$ ($\alpha=5\%$)

となり、年代に依って大きく異なる判定を齎す。即ち、検定の有意性が確認されたのは古代に於いてのみであることが分かる。表22に於ける実測度数の分布を見ると、1949（昭和24）年時点では未だ古代の分布の名残を留めているが、20世紀も後半になってくると古代の特徴は完全に失われていることが分かる。ここで注目すべきことは、古代に於ける郡名長の音列パターンは、先に表4で確認された日本語基礎語彙のそれとよく似た特徴をもっているという事実である。このことは、基礎語彙というものが、一般に当該言語固有の古い時代の特

表22：郡名データ（3音節）に於ける音韻パターンの推移。母音分布は、和名抄記載訓に対して、a：300，i：194，u：106，e：22，o：86，1949年データに対して、a：241，i：162，u：87，e：35，o：108，1997年データに対して、a：226，i：150，u：76，e：35，o：77，2010年データに対して、a：156，i：104，u：53，e：25，o：61。

(a) 古代→昭和24年

パターン	和名抄記載訓		1949年	
	実測度数	期待度数	実測度数	期待度数
ABC	62	76.79	74	79.09
AAB	65	45.06	50	38.37
ABA	44	45.06	39	38.37
ABB	33	45.06	27	38.37
AAA	32	24.03	21	16.81
合計	236	236.00	211	211.01

(b) 平成9年→平成22年

パターン	1997年		2010年	
	実測度数	期待度数	実測度数	期待度数
ABC	69	67.85	54	49.22
AAB	33	34.51	20	24.22
ABA	38	34.51	28	24.22
ABB	30	34.51	19	24.22
AAA	18	16.62	12	11.12
合計	188	188.00	133	133.00

徴を如実に反映したものであることと整合する。

4.5 編名、巻名、其他

是迄述べた2音節語に対する音列パターンの解析では、男子名の場合を除き全て検定の有意性が認められず、母音の配置が確率的に成されていることが明らかとなった。ここでは更に詳しくこのことを考察する為、標記の例を取り上げることにする。具体的には、六十干支名称、韓非著『韓非子』全55編の編名、及び道元(1200-1253)著『正法眼蔵』全72巻の巻名の中から音節数が2であるものを選び、これらに対して本手法を適用する⁽⁸⁾。各音列パターンへの分類結果を表23に示す。これより

六十干支に於いて、 $\chi^2=0.24 < 3.841$ ($\alpha=5\%$)

韓非子に於いて、 $\chi^2=0.31 < 3.841$ ($\alpha=5\%$)

正法眼蔵に於いて、 $\chi^2=0.51 < 3.841$ ($\alpha=5\%$)

となって矢張り検定の有意性が認められない。

4.6 名前ランキングのモーラ分析

最後に、名前の解析を通して明らかとなった音韻パターンの特異性を別の視点に基づいて考察する為、当該

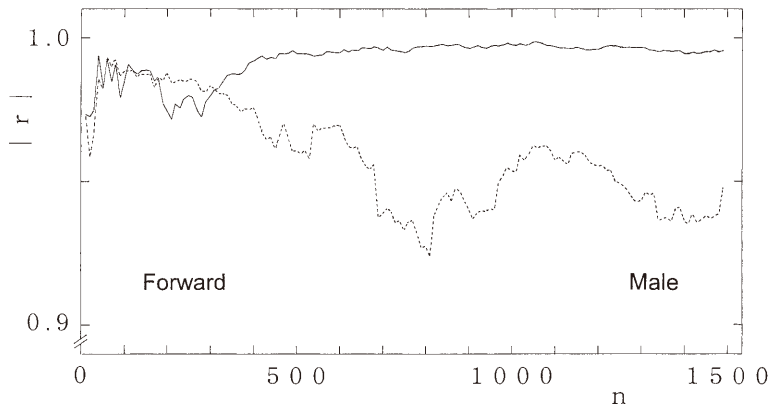
表23：六十干支，『韓非子』編名，『正法眼蔵』巻名のうち，2音節のもの音韻パターンの度数分布．母音分布は，六十干支に対して，a：1，i：43，u：6，e：0，o：22，『韓非子』に対して，a：12，i：9，u：9，e：4，o：18，『正法眼蔵』に対して，a：7，i：12，u：16，e：11，o：22．

パターン	六十干支		韓非子		正法眼蔵	
	実測度数	期待度数	実測度数	期待度数	実測度数	期待度数
AB	21	19.54	21	19.79	28	26.25
AA	15	16.46	5	6.21	6	7.75
合計	36	36.00	26	26.00	34	34.00

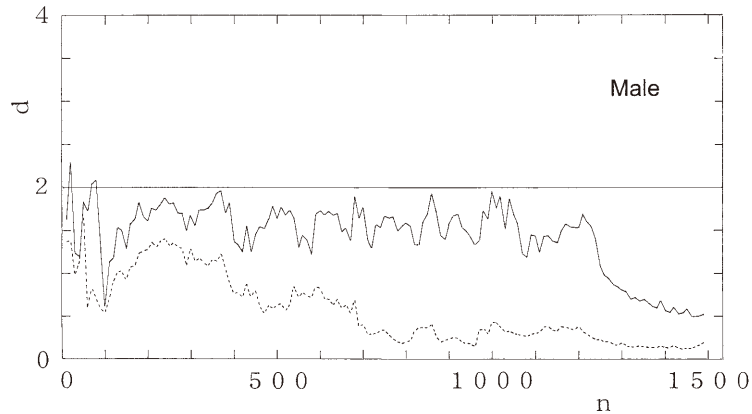
データのモーラ分布を考える．ここにモーラ (mora) とは拍とも言い，是迄扱ってきた音節が母音を中心とした音の纏まりであるのに対して，時間的に同じ長さを占める音の纏まりを指す．日本語に於いてモーラとは単語を仮名で表記した際の各々の文字に相当する (中島，2005：655)．従って，例えば「しんいちろう」は4音節6モーラ，「しょうこ」は2音節4モーラとなる．解析の手順を以下に記す．

- (1) 名前ランキングに表記されている仮名 (45種類) の度数分布を求め，度数の高いものから順位を付ける．但し，名前に濁点・半濁点が含まれる場合にはこれを取る (例：「が→か」，「ぼ→ほ」)．
- (2) 各仮名の累積度数を y ，順位を x とすると，45個の座標 (x_i, y_i) を得る．ここに i は1から45までの整数を表す．
- (3) 座標 (x_i, y_i) の散布図を互いに双対な二通りの方法で描く．一つは $(y_i, \log x_i)$ ，もう一つは $(\log y_i, x_i)$ である．ここに対数の底は10 (常用対数) とする．
- (4) 各散布図に於いて回帰直線を求める．ここで便宜上， y 対 $\log x$ ， $\log y$ 対 x の関係式を其々対数分布，指数分布と呼ぶ．各直線への適合度は，其々 y と $\log x$ ， $\log y$ と x の間の相関係数 r ($|r| \leq 1$) を用いて定量化する．尚，回帰分析の妥当性はダービン・ワトソン検定 (武藤，1995：208-209) に拠り判定する．

以上の手順に従って得られた仮名分布の解析結果を，男子名ランキングに対して図1に，女子名ランキングに対して図2に示す⁹⁾．両分布への適合度 $|r|$ 及びダービン・ワトソン比 d ($0 \leq d \leq 4$) の推移状態から，男子名データに対して対数分布 (実線) への高い適合性が認められる．ここに， $|r|$ の値が1に近く且つ d の値が2に近い程，当該分布への適合度が高いことに注意されたい．図1と類似した結果は，是迄に音階や文章の解析に於いて言及されてきた (早田，1998，2001)．取り分け，単調とも乱雑とも言えない，適度に変化を含む時系列データに対して，このような挙動を示す傾向にあることが是迄の経験から分かっている．これに対し，女子名データに対しては，非常に狭い領域，具体的には $750 < n < 800$ 付近でのみこの分布への適合性が見られるが，全体として両分布への適合性は認められない．性別に由るこのような質的な相違は，先に表9-11に於いて明らかになった事実と矛盾しない．以上を要するに，女子名の決定に於いては，長さや音列パターンのみならず，モーラに就いても歴史的・文化的制約を受けていることを示唆している．これとは対照的に，男子名の決定に於いては，長さや音列パターンと同様，モーラに就いての制約が弱く，モーラ同士が競合状態にあるものと推察される． $n=1000$ に於ける断面図を対数分布に対して図3に示す．ここに a ， b は其々 y 軸上の切片，傾きを表す．次に，名前データに於ける方向性の有無を探る為，図1，2のときは解析の順序を逆転し，ランキングの下位から上位へ累積度数を取った場合の結果を図4 (男子名) 並びに図5 (女子名) に示す．図1と図4の比較から明らかなように，男子名データの対数分布 (実線) に対して，顕著な反転非対称性が観測される．



(a)



(b)

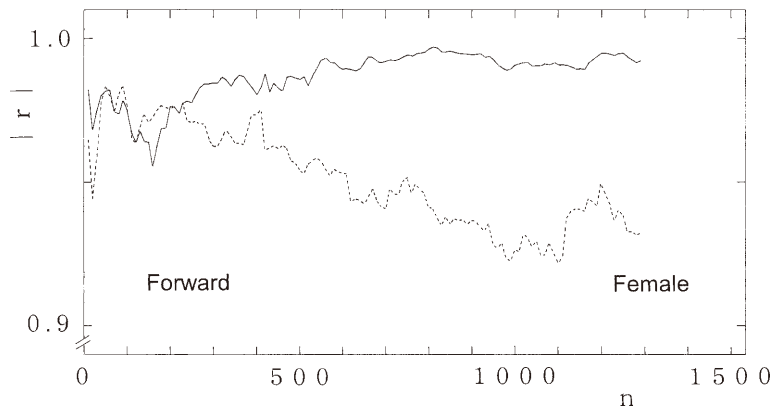
図1 男子名ランキングに於ける仮名(45音)の累積度数分布に対する解析結果。実線、破線は其々対数分布、指数分布への適合状態を表す。 $|r|$ は適合度、 d はダービン・ワトソン比、 n はデータの大きさでランキング表に於ける順位(上位から下位へ)と一致する。(a) $|r|$ の推移。(b) d の推移。

即ち、順方向(図1)の場合とは異なり、逆方向(図4)では d の値が略全領域で1を下回っており、又 $|r|$ の値も全般的に低下していることが分かる。このような順序反転に伴う非対称性は、音階や文章の解析で屢々見られる現象である(早田, 1998, 2001)。

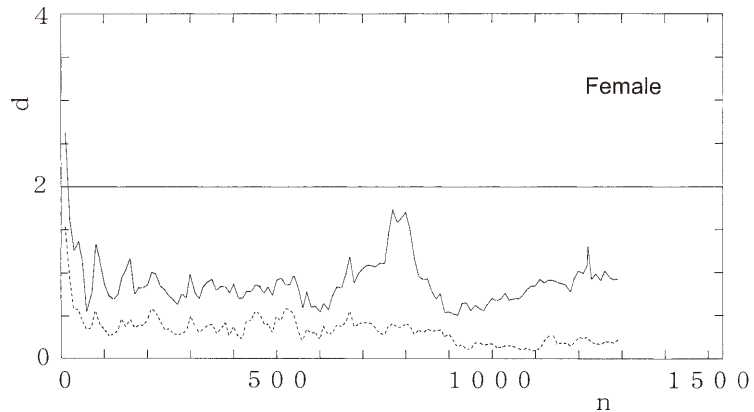
5 むすび

統計的手法に拠り、各種日本語語彙データに於ける音韻パターンの解析を行った結果を示した。データ規模は、苗字と名前に対して共に三千、地名に対して約二千である。ここに、名前に就いては男女別々に解析し、地名に就いては全市町村並びに郡を対象とした。更に比較の為、日本酒銘柄及び枕詞に対しても同様な分析を行った。本論文に因って明らかとなった主な事実を要約すると以下のようなになる。

- (1) 日本語基礎語彙(3音節)の母音配置を解析した所、五つの音列パターンのうちで、AAB型への偏りとABB型に対する回避が生じていることが分かった。
- (2) 苗字(3音節)に就いても(1)と同じ結論が導かれた。4音節語では、AABC, AABB型への集中と、ABCD,



(a)

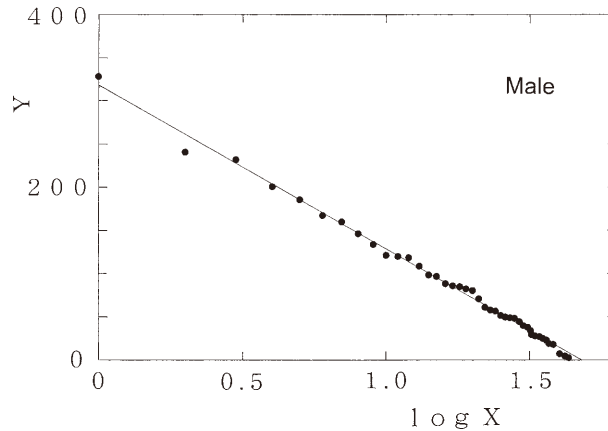


(b)

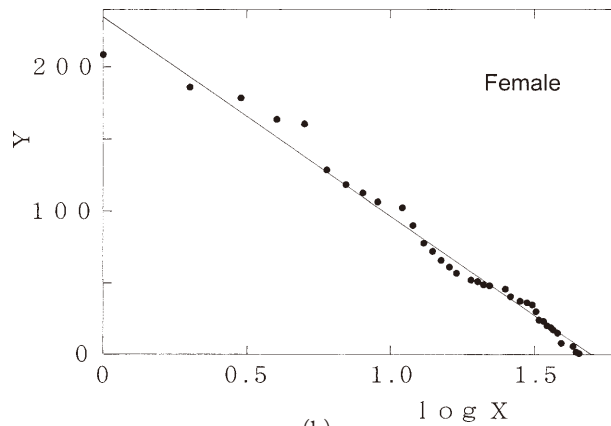
図2 女子名ランキングに於ける仮名(45音)の累積度数分布に対する解析結果。実線、破線は其々対数分布、指数分布への適合状態を表す。 $|r|$ は適合度、 d はダービン・ワトソン比、 n はデータの大きさでランキング表に於ける順位(上位から下位へ)と一致する。(a) $|r|$ の推移。(b) d の推移。

ABBC, ABAB型及びABBA型に対する回避が認められた。

- (3) 男子名(2音節)では、AB型とAA型の二つのパターンのうち、AB型への偏りが生じていることが分かった。また、3音節語、4音節語に対しても、其々ABC型、ABCD型への顕著な集中が見られた。以上を要するに、男子名に於いては、音節数に拘らず、近接する母音の間で同音回避状態が好まれていることが明らかとなった。
- (4) このような同音回避現象は、女子名(3音節)に於いても生じていることが分かった。但し、男子名とは異なり、(1)と(2)で言及したAAB型への偏りも見られた。更に、ABB型への注目すべき忌避傾向が確認された。
- (5) 現代地名(3音節)に就いては、市・町村・郡共に、音列パターンが確率法則に従っていると結論された。一方、古代郡名(同)に就いては、(1)及び(2)と同じ状態にあることが確認された。
- (6) 枕詞(助詞を除く)に対しては、語頭部に同音が集まった、AABC型とAAAB型の二つのパターンが非常に好まれることが示された。



(a)

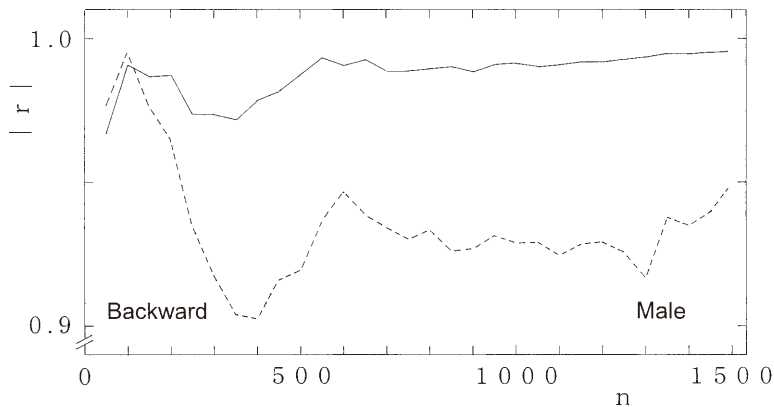


(b)

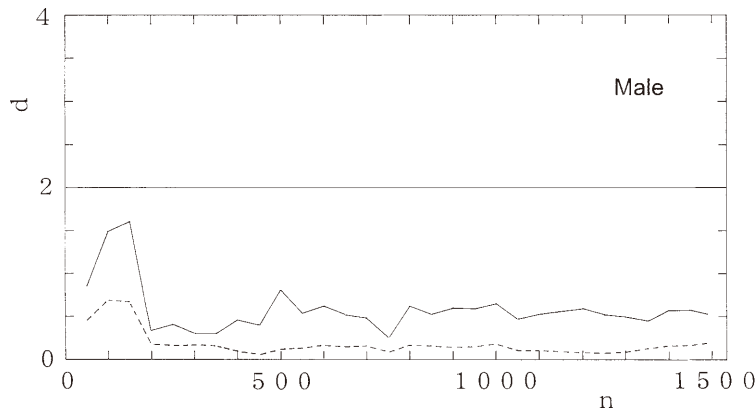
図3 $n=1000$ (順方向) に於ける対数分布への適合状態. y は仮名の累積度数, x は仮名の順位を表す. (a) 男子名ランキングに対して. 点の総数は41で左から {う, い, し, た, き, と, か, ゆ, ん, ろ, け, ひ, よ, す, り, つ, や, あ, こ, さ, ま, ち, お, み, く, の, る, て, ふ, な, も, そ, ら, せ, む, は, え, へ, れ, ほ, め, に, ね, わ, ん} に対応している. $|r|=0.9976$, $d=1.957 > 1.55$ (5%), $(a, b)=(317.75, -189.42)$. (b) 女子名ランキングに対して. 点の総数は36で左から {み, な, か, こ, り, あ, ゆ, き, い, さ, え, の, す, ま, し, つ, は, ほ, う, る, ん, お, と, も, ら, ち, や, よ, ひ, く, ね, れ, め, た, せ, ろ, ふ, わ, そ, む, に, け, て, ん, へ} に対応している. $|r|=0.9903$, $d=0.703 < 1.41$ (5%), $(a, b)=(234.54, -138.12)$.

(7) 日本語基礎語彙と比較する目的で, 主要な五つのロマンス諸語に就いて同様な解析を行った. フランス語に於いては, 母音字を二つ含む語及び三つ含む語で, ポルトガル語に於いてはこれを三つ含む語で, 近接する母音字間で同字回避状態にあることが分かった. ここで明らかとなった事実は, フランス詩法の文脈で言及されてきた事項 (鈴木, 2008: 181) と矛盾しない.

(8) 名前ランキングに対して仮名45音の累積度数分布を求め, これを順位付けして, 対数分布と指数分布への適合度を調べた. その結果, 男子名データが対数分布へ非常によく適合していることが明らかとなった. この事実は上述の(3)と矛盾しない. 男子名ランキングに於いてはこれに加え, 文章に対して屢々観測されたような顕著な反転非対称性が認められた.



(a)



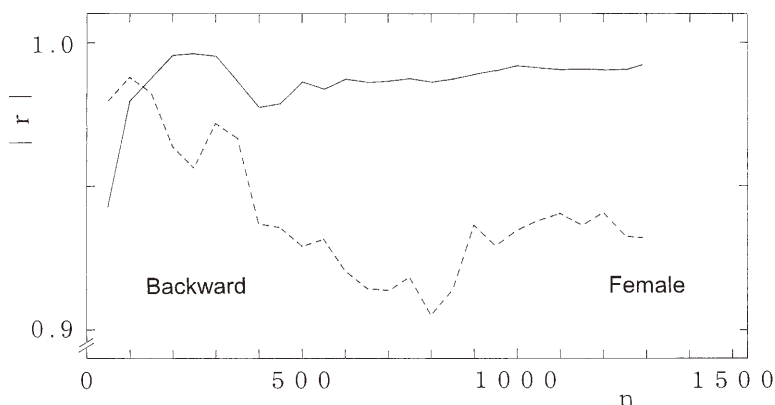
(b)

図4 男子名ランキングに於ける仮名の累積度数分布に対する解析結果. n の取り方が図1と逆向き(下位から上位へ)である他は図1と同じ. (a) $|r|$ の推移. (b) d の推移.

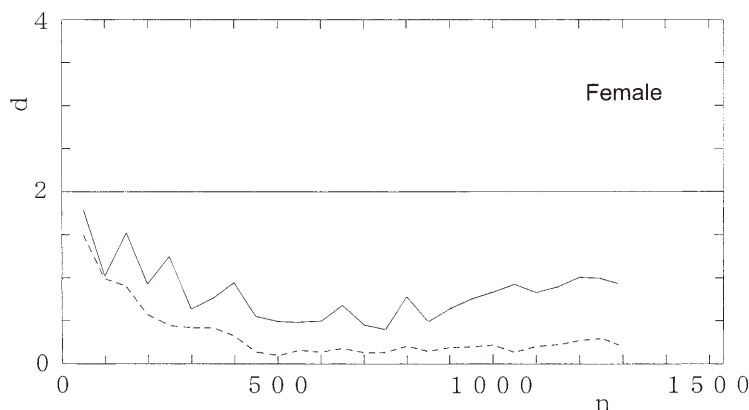
注

- (1) 後者の例として身近なものに「札幌郡広島町」→「北広島市」がある(既存の広島県広島市との同音衝突を避ける為). ジリエロン (Jules Gillieron, 1854-1926) の言語地理学に拠れば, 語彙一般の場合, このような同音衝突の回避現象は, 地理的に近接した地域間で生じるとされるが, 殊地名に関しては, 北海道と広島県の如く地理的に隔たった地点の間でも本現象が起こることが分かる. 所が近年 (2006年1月1日), 「伊達郡伊達町+梁川町+保原町+霊山町+月舘町」→「伊達市」(福島県) のように, この慣例を破る例が発生した. 本特例に拠り, 長年唯一の例外であった「府中市」(広島県, 東京都) に加え, 伊達市は北海道と福島県で併存することになった. 尚, 今日では, 嘗ての「江戸」(穢土と同音) →「東京」で見られたような不合理な理由で地名を変えることは禁止されている. 但し, 駅名に関しては, 札幌地下鉄南北線でこれと似た事案があったことは記憶に新しい. 旧「霊園前」→新「南平岸」と改名.
- (2) 異国と言っても矢張り西洋に偏っているようである. インド風やバルシャ風, アラビア風の名前があってもよさそうなのだが中々見当たらない. 嘗て森鷗外 (1862-1922) は自分の五人の子全てに洋風な名前を付けたことで知られている(森, 1993: 294-295). 息子三人には其々「於菟 (Otto)」「不律 (Fritz)」「類 (Louis)」, 娘二人には其々「茉莉 (Marie)」「杏奴 (Anne)」と命名した. このうち鷗外が最も期待を掛けていた長男の於菟 (1890-1967) は, 台北帝国大学を経て東邦大学医学部で解剖学の講座を担任した. 名前に対する鷗外の趣向は孫にも及び, 於菟の五人の息子は順に「真章 (Max)」「富 (Tom)」「礼於 (Leo)」「樊須 (Hans)」「常治 (George)」と命名された. このうち四男の樊須 (1928-2007) は北海道大学農学部で応用動物学の講座を担任した. 生前は江別市大麻園町に住んでいた.

所で, 本居宣長著「玉勝間」八の巻に, 「男の名にも某子といえる事」と題した記事がある. これに依ると「中昔よりこな



(a)



(b)

図5 女子名ランキングに於ける仮名の累積度数分布に対する解析結果。nの取り方が図2と逆向き(下位から上位へ)である他は図2と同じ。(a) $|r|$ の推移。(b) dの推移。

た、女名に某子といふこと、なべての例也、いにしへにもをりをり見えたり、さていにしへは、男の名にも、子といへる多し」とあり、古事記などを典拠に、口子(くちこ)、眞根子(まねこ)、浦嶋子(うらしまこ)といった多数の実例を引用している(本居, 1934: 320-321)。時代が下ると伴に、男子名に「子」の字を添える慣行は急速に廃れた。(同様な傾向は「磨呂(まる)」に就いても言える。)嘗て、虚子、秋桜子、誓子など俳人の雅名として、この字の使用が流行したことがあったが、音は訓読み「こ」ではなく音読み「し」となっている。現在では、男子名語尾音が「こ」であるものは、僅かに「ひこ」という形で名残を留めているに過ぎない。「ひこ」の表記としては、「彦」が広く用いられているが、「比古」という記載も稀に見られる。「比子」「牌子」「肥子」「被子」「婁子」「日子」などの表記は今では見たことがない。

- (3) 本論文ではロマンス諸語として五つの主要なものを解析の対象としたが、他のロマンス諸語、具体的にはカタルーニャ語、オック語、フランコ・プロバンス語、レト・ロマンス語、サルデーニャ語(町田, 2011)との比較は大変興味深い課題である。
- (4) 長名の例(稲, 2007)として面白いものを二、三挙げると、「春鶯囀(しゅんのうてん)のかもさる蔵(山梨県)」「酒屋八兵衛伊勢錦(三重県)」「酒家長春萬寿亀泉(高知県)。
- (5) 枕詞の総数は、基幹語から派生したものや単に詠ったものを数に入れるか否かで大きく食い違ってくる。主なものだけに限定すると350語程になる。江戸中期の国学者・歌人であった賀茂真淵(1697-1769)は、自身の枕詞研究書『冠辞考』の序文(井上, 1978: 1-5)に於いて次のように述べている。「此言上つ世の上より、中つ世のくだちに至りて、今に傳れるいろは三百まり、数は六百にもたりやしぬらん、」更に「凡冠辭をこれにとりし数は、三百四十餘あり、且その一つの中に、小別のいと多きあり、それを数へは、六百の有なんといふ也、」尚、この真淵の『冠辞考』に対して下された小林秀雄(1902-1983)

の論評は中々出色の出来栄である(小林, 1992: 224-231). そこでは、「冠辞は、勿論理論にも実用にも無関係な措辞だが、思い付きの贅語でもない。ひたすら言語の表現力を信ずる歌人の純粋な喜び、尋常な努力の産物である。」とある。枕詞の本質を突いた文である。

(6) 『冠辞考』(井上, 1978) 所収全350語の冠辞(枕詞)を調べた所、語尾に助詞(1音節)が付いているものの総数は152であった。これらに就いて表12と同様な分類を行うと、矢張り音列2) AABC, 8) AAAB並びに12) AABBへ顕著な集中が見られた。今、fとFを其々実測度数、期待度数とすると、AABC型に対して(f, F)=(24, 13.94), AAAB型に対して(f, F)=(12, 6.79), AABB型に対して(f, F)=(12, 5.97)であった。

(7) 5母音体系をもつ言語の数は全母音体系の中で最大で、684言語のうち221言語であると報告されている(松本, 2006: 166)。この値は全体の32%に相当し、6母音体系(第二位)の18%を大きく引き離している。5母音体系をもつ主な13の世界言語のうち、5母音中{a}の出現確率が最大なものは、スワヒリ語(47.13%)、サモア語(42.80%)、マルタ語(41.62%)、グルジア語(39.11%)、アルメニア語(36.16%)、現代ギリシア語(31.25%)、古典ギリシア語(31.04%)、アイヌ語(30.96%)、日本語(29.86%)と続く。ここに括弧内の数値は出現確率を表す。他の4言語のうち、5母音中の{a}の順位は、スペイン語(29.51%)で第二位、ロシア語(25.14%)、チェコ語(20.07%)、ラテン語(16.71%)で第三位である。尚、日本語固有の特徴として、他の言語と比べ{o}の使用率が際立って高いことが挙げられる。実際、上記13言語の中でこの値が最大で、使用率が28.25%と見積もられている(因みに第二位はロシア語の25.23%)。この値は13言語の平均値17.09%を大きく上回っている(松本, 2006: 32)。従って、次のような言い回しも日本語として何等違和感がない: 「男心のこの事を(otokogokoro no kono koto o)」

(8) 六十干支のうち、次の36干支がこれに相当する。

甲子, 乙丑, 戊辰, 己巳, 庚午, 辛未, 壬申, 癸酉, 戊寅, 己卯, 庚辰, 辛巳, 壬午, 癸未, 甲申, 戊子, 己丑, 庚寅, 辛卯, 壬辰, 癸巳, 甲午, 庚子, 辛丑, 癸卯, 甲辰, 戊申, 己酉, 壬子, 癸丑, 甲寅, 戊午, 己未, 庚申, 辛酉。

『韓非子』全55編のうち、以下の26編がこれに該当する。

存韓, 難言, 主道, 有度, 十過, 孤憤, 和氏, 亡微, 三守, 南面, 喻老, 觀行, 安危, 守道, 用人, 功名, 難二, 難三, 難四, 問弁, 問田, 詭使, 五蠹, 忠孝, 人主, 心度。

『正法眼藏』全72巻のうち、次の34巻がこれに対応する。

仏性, 空華, 光明, 行持, 恁麼, 觀音, 古鏡, 有時, 授記, 全機, 都機, 画餅, 看經, 伝衣, 道得, 仏教, 神通, 春秋, 葛藤, 嗣書, 仏経, 法性, 洗面, 面授, 仏祖, 洗淨, 十方, 遍參, 家常, 龍吟, 虚空, 鉢孟, 安居, 出家。

(9) 男子名全体(n=1487), 女子名全体(n=1285)に対する仮名の内訳は以下の通り。小括弧, 大括弧内の数字は其々45音中の順位, 累積度数を表す。

男子名: 上位5位→(1)う [455], (2)し [364], (3)い [328], (4)た [309], (5)と [268]

下位5位→(41)ほ [14], (42)に [12], (43)め [9], (44)わ [4], (45)ぬ [0]

女子名: 上位5位→(1)み [271], (2)な [236], (3)か [221], (4)り [207], (5)こ [198]

下位5位→(41)け [11], (41)に [11], (43)て [10], (44)へ [3], (45)ぬ [2]

参考文献

- 群馬県立大学北川研究室: 編 (2009).
<http://homepage1.nifty.com/k-kitagawa/data/gunmeil.html>
早田和弥 (1998). コンドラトフ人工賛美歌の音楽性分析. 『社会情報』, Vol.8, No. 1, 109-114.
早田和弥 (2001). 非情報論的手法による文学テキストの解析/合成と可視化. 『第6回日本社会情報学会大会発表要旨集』, 52-55.
Hayata, K. (2006). Statistical prosody: Rhyming pattern selection in Japanese short poetry. *Forma*, Vol. 21, 259-273.
稲 保幸 (2007). 『日本酒15,706種』, 誠文堂新光社.
井上 豊: 編 (1978). 『賀茂真淵全集第八巻』, 統群書類従完成会.
伊藤太吾 (1994). 『ロマンス言語学入門』, 大阪外国語大学学術出版委員会.
小林秀雄 (1992). 『本居宣長(上)』, 新潮社.
町田 健 (2011). 『ロマンス語入門』, 三省堂.
松本克己 (2006). 『世界言語への視座』, 三省堂.
森 於菟 (1993). 『父親としての森鷗外』, 筑摩書房.
本居宣長: 著, 村岡典嗣: 校訂 (1934). 『玉勝間(上)』, 岩波書店.
村山忠重 (2000). 『日本の苗字ベスト10000』, 新人物往来社.
武藤眞介 (1995). 『統計解析ハンドブック』, 朝倉書店.
内藤弘作 (2008). 『枕詞便覧』, 早稲田出版.
中島平三: 編 (2005). 『言語の事典』, 朝倉書店.

日本地方行政研究会：編（1949）.『全国市町村便覧』, 全国教育図書.
西沢教夫, 牧野恭仁雄（2001）.『日本の「なまえ」ベストランキング』, 新人物往来社.
鈴木信太郎（2008）.『フランス詩法 上巻』, 白水社.
矢野恒太記念会：編（1996）.『データでみる県勢 1997年版』, 国勢社.
矢野恒太記念会：編（2007）.『データでみる県勢 2008年版』, 矢野恒太記念会.
矢野恒太記念会：編（2009）.『データでみる県勢 2010年版』, 矢野恒太記念会.
安本美典（1995）.『言語の科学』, 朝倉書店.
郵政省郵務局：編（1999）.『新郵便番号簿 平成11年版』, 郵便局.

Selection Rules for Phonemic Patterns in Japanese Names concerning Families, Persons, and Places

by

Kazuya HAYATA *

Abstract

Phonemic patterns of Japanese name data for families, persons, and places are analyzed statistically, where, in addition to cities, towns, and villages, ancient and modern counties are contained in places. For both families and persons the scale of the data attains three thousand, whereas for places it attains about two thousand. In order to make a comparison, brands of *sake* as well as 'pillow words' are also considered. Analyzed results of the vowel orderings in the entire words show that, in contrast to the stochastic law being found in the modern place names, for family as well as personal names, their phonemic patterns are not determined stochastically. In particular, the patterns of personal names differ significantly from those confirmed in the basic vocabulary of Japanese. Specifically, irrespective of sex, avoidance is seen for the identical vowels in neighboring syllables.

Keywords: Vocabulary Statistics, Name Analysis, Toponymy, Phonemic Law, Pillow Words

*Faculty of Socio-Informatics, Sapporo Gakuin University, hayata@sgu.ac.jp