

# ケータイメールにおける顔文字と記号の 出現頻度とその関係

—ケータイメールコーパスの紹介とともに—

加藤 安彦

## はじめに

私たちはいまや携帯電話の洪水の中にいる。道を渡りながら、公園で寝そべりながら、近くか遠くかあるいは海外か、ともあれ普通にはケータイと呼ばれているこの手のひらサイズの「文明の利器」を通じて、誰かと「今現在」を共有している。

ぼくは時々、世界中の電話という電話は、みんな母親という女性たちのお膝の上なんかについているのじゃないかと思うことがある。特に女友達にかける時なんかがそうで、どういうわけか、必ず「ママ」が出てくるのだ。

庄司薫の芥川賞受賞作、「赤頭巾ちゃん気をつけて」の冒頭部分である。もう四十年近く前に世に出た小説ながら文体もまだ今風といってもよさそうだし、現代の若者にも十分受け入れてもらえそうに思う。しかし、そうであるにもかかわらず、この冒頭部分で現代の若者たちはつまづいてしまう。理解するのに少々困難な面がある。電話がみんな母親という女性たちの膝の上にあるからといって、それがどうしたというのだろう、ということになるにちがいない。

この庄司薫の「赤頭巾ちゃん気をつけて」初版は、昭和四十四年刊行であるが、その一年前、昭和四十三年に、母親という女性たちのお膝の上なんかについていると書かれた固定電話の施設数が、日本全体で1,000万件を超えている。その後も固定電話の施設数は順調に数を伸ばしていき、平成六年には6,000万件を超えたが、平成九年十一月の6,322万件をピークに施設数は減少傾向に転じる<sup>註1</sup>。

## 2 専修国文 第81号

代わって、同年同月における携帯電話の契約台数は2,766万台であったが、携帯電話の契約数は固定電話とは反対に増加しつづけ、平成十七年十一月現在で8,968万台、つまり、約9,000万台で固定電話施設数の約1.5倍の台数が普及している<sup>註2</sup>。

日本全体の世帯数は平成十七年の国勢調査において4,953万世帯であると報告されている<sup>註3</sup>。会社や官公庁など機関・組織での固定電話契約、携帯電話契約もあり、少々乱暴ではあるが、単純計算をすれば1世帯に1.2台の固定電話に加え、1.8台の携帯電話があることになる。

さて、日本語学専攻の私が担当するゼミでは、ゼミ生の協力を得て、この新たなコミュニケーション手段である、携帯電話間でのメール、携帯メールの収集を行っている。まだ黎明期にあるといえるこの携帯メールのやりとりを電子化して保存し、将来的には携帯メールコーパスとしてさまざまな研究利用に耐える資料としていきたいと考えている。

携帯電話のメール保存機能は、せいぜい200通程度であり、時間経過と共に古いメールから次々と消滅していつてしまうため、この時期のこうした言語資料が永遠に失われてしまうことを少しでも防ぎ、黎明期の携帯メールを将来に互って残していきたいという気持ちが最大の動機である。また、協力してくれるゼミ生たちにとっても、お互いに携帯メールデータを提供しあい、同時にゼミ卒業生たちのデータをも資料とすることができるため、個人では収集不可能な、充分な量の携帯メールを卒業論文の考察対象とできるメリットもある。

さて、「専修国文」においては初めてこのデータを紹介するにあたり、まずはその概要を説明し、データの主たる提供者である若年層が携帯メール上のパラ言語表現である顔文字によって、どのような「気の遣い方」を見せるかについて述べることにする。

### 1. データ概要

本稿では、先述のとおり、ゼミで2003年度から入力を開始した収集データのうち、データの入力ミスや付加情報を付与し終えた、2004年度卒業生（ゼミ第一期卒業生）のデータを資料として使用する<sup>註4</sup>。因みに、本稿では、現在、

一般的である、「メール」という表記を使わずに「メイル」という表記を採用している。これは、「メーキャップ」が「メイクアップ」に、「ヘアメイク」が「ヘアメイク」に変化したこと、「メーンストリート」、「メーンイベント」が「メインストリート」、「メインイベント」に変化していることから、いずれ「メール」も「メイル」に取って代わられることを予想し、少し先取りする気持ちで「メイル」という表記を採用している。

### 【携帯メール入力マニュアル】

最初に、データ入力についてであるが、入力のためのマニュアルを作成しており、それに従ってデータ入力を行っている。このデータ入力はテキスト形式のデータであり、それを元にしてさまざまな情報を付加し、最終段階では EXCEL 形式のデータに加工してある。以下の「携帯メール入力マニュアル」によって、データにどのような情報を盛り込み、どのような構造を持たせているか、また、携帯メールに出現する、従来 of 書きことばには見られなかった種類の文字をどのように処理しているか示すことにする。

### 《入力項目》

1 通のメールのデータは、以下の 3 つの部分からなっている。

3 つの部分とは、

- 1：データ管理行
- 2：メール題名行
- 3：メール本文

である。

それぞれの入力項目については、下記のとおり。

#### ○1 行目 〈データ管理行〉

以下の 6 項目について記述する

#### ①自分の携帯電話の会社名（記号は大文字全角）

D：docomo

V：Vodafone/SoftBank

4 専修国文 第81号

E : au/TU-KA

P : PC

H : PHS

W : WILLCOM

②送受信の別（矢印の向き）

送信：自身→相手

受信：自身←相手

③相手の携帯電話の会社名（①の記号に準ずる）

④メールをやりとりした相手のID番号

⑤送受信年月日（半角数字）—（YYYYMMDDの形で）

⑥送受信時刻（半角数字）—（HHMMの形で）

※各項目の間には半角カンマを入れる

※不特定多数の相手の場合は、①と③にDIV（全角記号）を用いる。

※メイリングリストの場合は、①と③にML（全角記号）を用いる。

○2行目〈メール題名行〉

以下の1項目について記述する。

①題名：（全角コロン）

例）題名：おはよう！

※無題の時には「題名：（空白なし改行）」とする。

※題名に「Re：（Re>）」がついているもの、または「Re：（Re>）」の後の文が変わっていないときは、無題の時と同様とする。

※「Fw：（すべて半角）」がついているとき（転送のとき）はそのまま入力する。

※Vodafone/SoftBankの“Message from Skymail”も無題の時と同様とする。

※2文以上の場合でも1文で入力する。

○3行目～〈メール本文〉

- ①メール本文（一文一行の体裁にする。改行のしかたは下記規則による。）

[本文入力の際の規則・注意点]

〈一文一行の体裁〉

- ・携帯メール本文は、一文一行の体裁とする。

一文とは下記の規則による。

- ①以下のものを原則として文の終端記号と考える

句点・絵文字・顔文字・記号

- ②絵文字・顔文字・記号が文中の語として捉えられる場合はそれを終端記号とせず、改行はしない

- ③ただし、一行の文章があまりにも長すぎる場合は、読点もしくは意味の区切れで改行としてよい

- ④文の途中でEnterがかかっている場合は、原文に忠実にそこで改行して一文とする

- ⑤句点と見なされる空白はそれを終端記号としてそこで改行とする

- ⑥上記以外の空白（行頭、読点代わりなど）はそのまま入力して、終端記号とはしない

- ・メールとメールの間は一行あける。

- ・メールの本文に誤字・脱字等、誤りがあってもそれを忠実に入力、再現する

〈絵文字・顔文字の入力〉

- ・絵文字・顔文字については下記規則によって入力する

- ①絵文字、顔文字は「％（）」と「」％」で絵文字・顔文字の別がわかるように「絵」「顔」の文字を付与してくる

- ②絵文字は、パソコン上で共通データとして表示できないため、携帯電話各社のインターネット上のホームページにあるユニコード表等によって番号

を入力し、それに代える

- ③顔文字は、半角文字は半角文字、全角文字は全角文字を用いて忠実に再現して入力する
- ④絵文字・顔文字をくくる括弧、%%、ユニコードは全角文字を用いる
- ⑤=（いわゆるゲタ）が本文中に出現した場合は、その記号=のままを入力する（この=記号は「げた」と入力して変換するもの）
- ⑥表情を表しているものは顔文字とみなす
- ⑦「^^」等は顔文字として扱う

#### 絵文字・顔文字入力例

絵文字→%%絵（絵文字のユニコード）%%

顔文字→%%顔（( \_ ) /）%%

顔文字→%%顔（ ^ ^ ）%%

顔文字→%%顔（ | \_ . ）%%（壁から半分顔が覗いているもの）

#### 〈記号の入力〉

- ①！、？、♪、☆、★などのパソコンで変換、入力可能なものについてはそのまま入力する
- ②「↑みたいに～」や「下記参照↓」などの矢印は「%%（↑）と（↓）%%」でくくらない

#### 〈「形」という分類の入力〉

- ①絵文字・顔文字・記号に準ずる携帯メール表現として「形」という分類を設ける
- ②「；」（汗を意味するセミコロン）、「orz」「挫折、自己嫌悪、絶望感などを表し、アルファベットもしくは記号の組み合わせが違いつづばった状態の人間を表すもの）、「ノシ」（カタカナの「ノ」と「シ」の半角文字の組み合わせで、ダッシュして駆け出した様子などを表す）など、顔文字ではないが、意味を持たされた記号の組み合わせに対しては、すべて「%%

「」と「」%%」でくり、「形」の文字を付与しておく

- ③「|\_|」は顔文字（|は壁で、顔が半分覗いている状態）として扱うが、「|\_」(壁の陰にすっと隠れたという意味の動作を表す)は表情が無いので「形」として扱う
- ④「超嬉しい~↑↑」や「失敗しちゃった↓↓」等に現れる、上向き及び下向きの矢印は、形の分類として扱い、それぞれ「%%形(↑↑)%%」、「%%形(↓↓)%%」と入力する
- ⑤「; ;」は、文脈上「汗」として使われている場合（汗が頬を滴り落ちていた様子を表す）は形に分類し、%%形(; ;)%%と入力し、泣き顔として使われている場合（両目から涙がこぼれている様子を表す）は、%%顔(; ;)%%と入力して扱いを分ける

#### 形の入力例

%%形(; ;)%%  
%%形(orz)%%  
%%形(ノシ)%%

#### 〈固有名詞の入力〉

- ①固有名詞は、プライバシーに関わるようなものはすべて抽象名詞に置き換えて括弧でくくる
  - ・会社名、部署名→〈組織名〉
  - ・学校名
    - 小学校→〈小学校名〉
    - 中学校→〈中学校名〉
    - 高校→〈高校名〉
    - 短期大学→〈短大名〉
    - 大学→〈大学名〉
    - 大学院→〈大学院名〉
    - 専門学校→〈専門学校名〉

- ②人名は、文中で「〈 〉」とでくり、〈人名〉とする
- ③人名に「ちゃん」「さん」「くん」などがついている場合は、下記のように扱う

- 〇〇ちゃん→ 〈〈人名〉 ちゃん〉、
- 〇〇さん→ 〈〈人名〉 さん〉
- 〇〇くん→ 〈〈人名〉 くん〉
- 〇〇ポン→ 〈〈人名〉 ポン〉
- 〇〇ッチョ→ 〈〈人名〉 ッチョ〉
- 〇〇ッチ→ 〈〈人名〉 ッチ〉
- 〇〇氏→ 〈〈人名〉 氏〉
- 〇〇先生→ 〈〈人名〉 先生〉

これらは一括置換せずにその都度入力とする

- ②商品名は括弧でくくなくてもよい
- ③有名人は括弧でくくなくてもよい
- ④文意に関わって、意味が通らなくなるようであれば、大学に関する部署名、組織名、地名などはそのまま入力する

#### 〈電話番号等の扱い〉

- ①携帯電話、据付電話、FAXの電話番号などは、括弧でくり、〈電話番号〉と入力する
- ②携帯電話、パソコン等のアドレスは、括弧でくり、〈アドレス〉と入力する
- ③その他、プライバシーに関することは、個々人の判断に任せ、①、②に準じた処理をするが、本人が公表しても構わないと判断したものはそのまま入力

#### 〈住所・地名の扱い〉

- ①住所は、括弧でくり、〈住所〉と入力する
- ②地名は 〈住所〉とせず、そのまま入力する



例) 新宿に6時ね!

- ③ただし、地名の入力によってプライバシーに関わる、第三者によって個人を特定できてしまうような場合は、これを括弧でくくり、〈地名〉としておく
- ④また、駅名、バス停名、団地名等々も上記に準じて、括弧によってくくり、抽象的な名詞の入力をもってその代用とする

〈パラ言語的な文字表現の扱い〉

- ①「(」と「)」とによっくられた、「笑」や「泣」といった、文字本来の意味によるパラ言語的な表現とそれに付随する部分に対しては、顔文字同様に「%% (」と「)」 %%」とによっくくることとする
- ②絵文字、顔文字、形とは異なり、特に文字を付しておくことはしない

パラ言語的な表現の入力例

(笑) → %% ((笑)) %%

笑 → %% (笑) %%

・笑 → %% (・笑) %%

(w → %% ((w)) %%

/w → %% (/w) %%

(泣笑) → %% ((泣笑)) %%

(苦笑) → %% ((苦笑)) %%

(凹) → %% ((凹)) %% (気持ちが凹んだという意味を表している)

〈入力例〉

\*受信メール (送信メールは矢印が逆になる)

docomo の携帯に au の携帯から、「ごめんね」というタイトルのメールを、2003年7月7日18時23分に受信した場合

1行目

○○○,19820123,神奈川,女,D,←,E,△△△,20030707,1823

2行目

題名：

3行目～

いいえいえ！

お安い御用だよ%%顔 ((o^o^o)) %%

授業のあとになっちゃうと思うんだけどいいかな??

ごめんねっ%%顔 (<\_<) %%

※ゼミメンバーとのメールのやり取りは、受信メールだけを入力すること

以上が入力マニュアルである。このマニュアルに従って入力されたデータには、さらに送受信相手との親しさの度合いである「親密度」、相手が学内の友人であるか、学外の友人であるかの別といった情報を付与して、相手との関係がコミュニケーションに与える影響も考察できるようにしている。なお、親密度は三段階に分け、深刻な内容の話をすることのできる相手、家族や親友と呼べる関係の友人を「親密度1」、「親密度1」に比べると親しさは薄れるが、比較的行動を共にすることの多い仲のよい友人などを「親密度2」とし、顔を知っていて、ことばも交わす普通の友人を「親密度3」としてある。これは必ずしも厳密な分類ではないが、例えば顔文字を使用する頻度を分析すると、親密度による傾向が見てとれるということから有意な情報であるといつてよい。

2005年度卒業生のデータは、総数約6,300通余、メール本文の総行数が16,000行強からなる。ただし、ゼミの卒業生は女性だけであったため、男性同士のメールがこの年度にはない。

#### 【2004年度データ内訳】

##### 〈送受信相手の性別〉

データの送受信相手の内訳は、ゼミ卒業生が女性だけであるため、女性から女性、女性から男性、男性から女性というパターンとなっており、そのメール数は下記の通りである。「性別不特定」としてあるものは、携帯の電話番号が

変更になった場合や、メールアドレスが変更になった場合など、性別に関わらず友人相手に一斉に送信されるメールなどを指す。

表1 送受信相手性別別メール数

2004年度	送受信相手性別					合計
	女→女	女→男	男→女	男→男	性別不特定	
メール数	3715	1052	1584	0	11	6362

さらに、送受信相手の性別とデータ全体におけるメール本文の総行数を示しておく。

表2 送受信相手性別 1メールあたりの行数

	送受信相手性別			
	女→女	女→男	男→女	合計
行数	10356	2366	4007	16729

〈送受信相手の親密度〉

次に親密度から見たデータの内訳を示す。この数値には、性別の確定できない不特定多数の相手に送信した、あるいは不特定多数の相手の一人として受信したデータは、親密度・性別が固定ではないため、ここに含めない。また、男女の別も考慮していない。

表3 送受信相手との親密度によるメール数内訳

	親密度			合計
	1	2	3	
メール数	3730	2109	512	6351

## 2. データ分析の試み：顔文字と記号類の関係

ゼミでのデータは、既に4年度分のデータが集まっており、さらにまだ今後も収集は続ける予定である。

ここでは、今回データ整備を終えた2004年度分のデータについて、総数は少

ないが、そこから見てとれる傾向について考察を行う。入力マニュアルにあるとおり、顔文字、絵文字をデータ上では「%% (」と「) %%」とで囲み、別立てとしており、また、男女の別や親密度といった情報も付与している。そこで、それらの情報付与が有意義であることを検証的に示しておきたい。

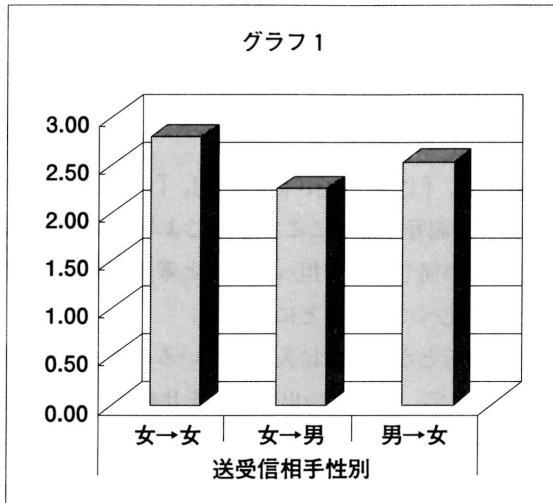
まず、今回のデータでは、男性から男性に出したメールは1通も存在しないが、先に示した表1、表2で示してある送信者の性別、送信相手の性別からなにか傾向があるかどうかを見てみることにする。表1、表2から、送信者と送信相手の性別によって、以下のような1メールあたりの行数が得られる（行数をメール数で割ったもの）。

表4 送信者・送信相手の性別にみた1メールあたりの行数

	送受信相手性別		
	女→女	女→男	男→女
行数/メール数	2.79	2.25	2.53

これをグラフにしたものが、次のグラフ1になる。

これを見ると、女性同士でのメールが最も長く、女性が男性に出したメールが最も短いという結果になっている。女性から男性より、男性から女性へのメールの方が行数が多めであることがわかる。ただし、1行あたりの文字数はカウントしていないので、どのような文字種で1行が成り立っているか、音節数がどれほどであるか、といったことはわからない。大体の傾向として上記のようなことが言えるということである。



この傾向が他の年度のデータにおいても続くものかどうか、さらに調査を進める必要があるが、男性の方が女性よりも短い文で改行してしまう傾向がある可能性がある。

次に、顔文字と記号の出現頻度と親密度とを絡めて何らかの傾向があるか探ってみる。ここでいう記号類とは、「☆」「★」「♪」といったものを指す。顔文字と記号類とは同じような文脈において出現するのではないかと実際のメールのデータをチェックしながら感じたことであったが、それを実際に数値として捉えてみることにする。

まず、同じような文脈で使われるという点についてであるが、顔文字と記号類の実例を下記に挙げてみる。

〈顔文字類〉

「アドレス変えました(\*o\*)v」

「アドレスが変わりました(>\_<)」

「アドレス変えたの(^-^)」

〈記号類〉

「アドレス変えました☆」

「ハロー♪アドレス変更しましたッ」

以上は簡単な例だが、「(^o^)」v、「( )\_<」、「(-)」といった顔文字と、「☆」や「♪」といった記号類とがここでは同じような文脈で用いられているのがわかる。では、この同じ役割を担っていると考えられる記号類と顔文字に違いがないのかどうか比べてみることにしたい。

データでは、1文で1行となるように入力しているので、各親密度全体の行数（文の数と一致）と顔文字、記号類の出現行数を比較すれば、親密度別に顔文字、記号類それぞれの出現比率がわかる。親密度1の相手とのやり取りは9900文、親密度2は5394文、親密度3は1435文であるので、顔文字類の出現比率と記号類の出現比率は以下ようになる（1文に複数回出現しても出現は1文としてカウントしてある）。

〈顔文字類〉

親密度1 1914文／9900文

親密度2 718文／5394文

親密度3 225文／1435文

これを出現比率に直してみると以下のようなになる。

親密度1 19.33%

親密度2 13.31%

親密度3 15.68%

〈記号類〉

親密度1 676文／9900文

親密度2 502文／5394文

親密度3 92文／1435文

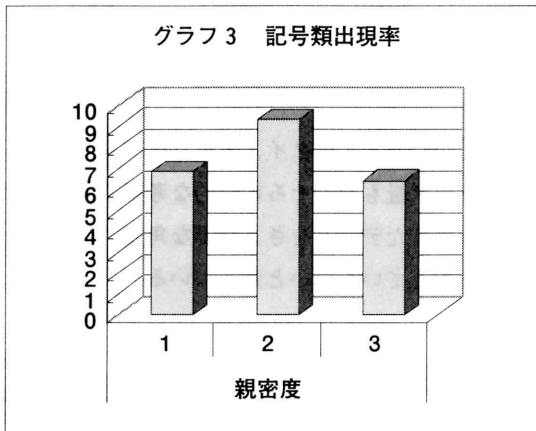
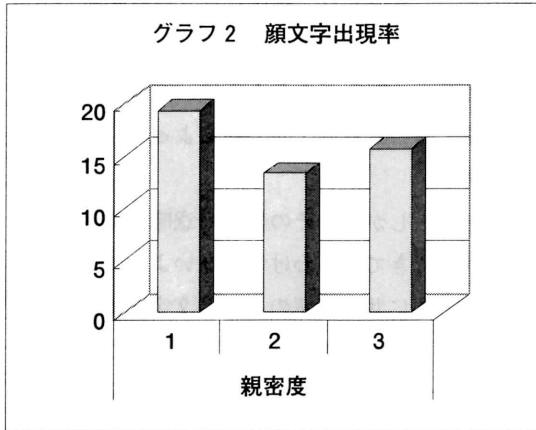
これを出現比率に直してみると以下のようなになる。

親密度1 6.83%

親密度2 9.31%

親密度3 6.41%

これをグラフにすると以下のグラフ2、グラフ3になる。



これら2つのグラフを見てわかることは、記号類の使用が特に親しくもなく疎遠でもない、という相手に対してよく使われるということ、それから、顔文字は親密度の最も近い相手と遠い相手によく使う、ということである。

これは、親密度1の親しい友人には顔文字は自分の気持ちを表現するのに有効なバラ言語的な表現として使用し、親密度3の普通の友人にはとりあえず笑

った顔をしたものを送ることで「あなたに対して悪意のないメールを書いているんです」という自らの態度を表明する役割を担わせていると考えられる。親密度2という、特に近くも遠くもない相手には「言外の言」ともいうべきパラ言語的な表現を使うにはやや相手が離れた位置におり、ただ笑った顔を送っておけばいいとするには近い存在であるために、記号類を他の親密度に比べるとより多く使っていると考えられる。記号類によって、にぎやかな雰囲気 of メールにしているのである。

顔文字には表情がある。しかし、その表情の意味するところはメールを送受信する相手と完全に共有できていないわけではないようである。そのことからすると、うっかり相手との間に共通理解のない顔文字を送ってしまって、意味のとり間違いが起ってしまうと、場合によっては顔文字が意図しない皮肉な意味合いを持ってしまったりすることになる。そのことで相手に不快感を与えるよりは、ひとまず記号類を使っているとも考えられる。グラフ2とグラフ3がお互いの補完的な様子からもそれは窺えそうである。

おわりに

今回はゼミで構築しつつあるケータイメールコーパスの紹介に重きを置いて書き進めてきたが、データ量もそろそろ本格的な考察に耐えるだけの量にはなってきたので、今後またデータをさまざまな角度、観点から分析・考察してみ、その結果を報告していきたいと考えている。

註

0 本稿はそのデータ作成及び研究考察の一部を、平成17年度日本語日本文学会研究助成金に拠っているものである。

1 [http://www.ntt-east.co.jp/fromn-east/200411\\_4.html](http://www.ntt-east.co.jp/fromn-east/200411_4.html)の【資料4 固定電話施設数及び施設設置負担金受入額の推移】による。

2 ただし、PHSは除く。

「社団法人電気通信事業者協会」ホームページ<http://www.tca.or.jp/>のデータベースによる。



- 3 「総務省 統計局・政策統括官（統計基準担当）・統計研修所」ホームページの国勢調査のページ「平成17年国勢調査全国・都道府県・市区町村別人口（要計表による人口）結果の概要」の「IV世帯」のページから。  
<http://www.stat.go.jp/data/kokusei/2005/youkei/04.htm>
- 4 2004年度卒業生データから始まり、現在に至るまでに収集された携帯メールデータは、すべて最終的に加藤が全体を取りまとめた後、加工を行ってゼミ生に提供しているものである。本稿においてもデータに関する不備等の責は加藤が負うものである。また、これらのデータはすべて提供してくれたゼミ生によって、その送受信相手全員にデータ提供と加工されたデータ利用についての承諾を受けているものである。本稿では、2004年度卒業生の、高橋まり子、曾我紗代子、鈴木麻紀子、小松原里依の4名のみさんのデータを使用した。記して感謝申し上げる。