

図1 サンプルスコアの年齢差 第1～4軸

し、若い世代で大きなマイナス値を取る。(後掲図4によれば、北半ではほぼ全年齢層に、南半では中間世代に多く、若い世代ではマイナス値になる。鶴岡付近がマイナス値なのは、鶴岡付近の若い世代が使う語形に結び付く。)

点線第4軸の年齢差はやや大きく、若い世代でさらに大きくなる。(後掲図5によれば、北半がプラス、かつ全域の若い世代がプラスで、今勢力を盛り返して使われていると見られる。)

まとめると、140年間隔のデータでは、年齢差が大きく働く。江戸時代に成立した方言集の、明治初期

生まれから平成期生まれまでの年齢差=見かけ時間 (apparent time) なので、近代語の普及、戦後の共通語化などが強く働いたと、読み取れる。従来のグロットグラム調査では、域内に方言差のある項目を選ぶことが多かったから、当然ながら地域差が目立った。『浜萩』の調査項目は江戸と庄内の違いに着目した方言集によるものなので、別の(むしろ無色の客観的な)基準で、選ばれたものである。250年の間にことばがどう変わるかを示す貴重な情報と言える。語彙統計学などの成果との突合せが可能である。

2.1 多重対応分析の個人ごとの値 第1軸

以下では4軸それぞれのサンプルスコアをグロットグラムに示す。図2に多重対応分析のサンプルスコアを、単純化グロットグラムによって地域差・年齢差を個人ごとに示す。縦軸は生年を示し、若い世代が下である。横軸は鶴岡からの徒歩距離を示す。地域を南北に分け、左側が南端鼠ヶ関、右側が北端吹浦と飛島である。値を以下の4個の区間に分けて、記号を与えた。

「 -0.5 0 0.5 」
「 ○ 。 # ■ 」

図2ではプラスとマイナスが大きく年齢層によって分けられる。その境界はほぼ1935年生まれで、第1次調査と第2次調査の境にあたるが、調査次の境界と一致するわけではない。境界付近には0.5未満の記号「#」。「。」が多いので、調査次とずれた時期にことばの大変

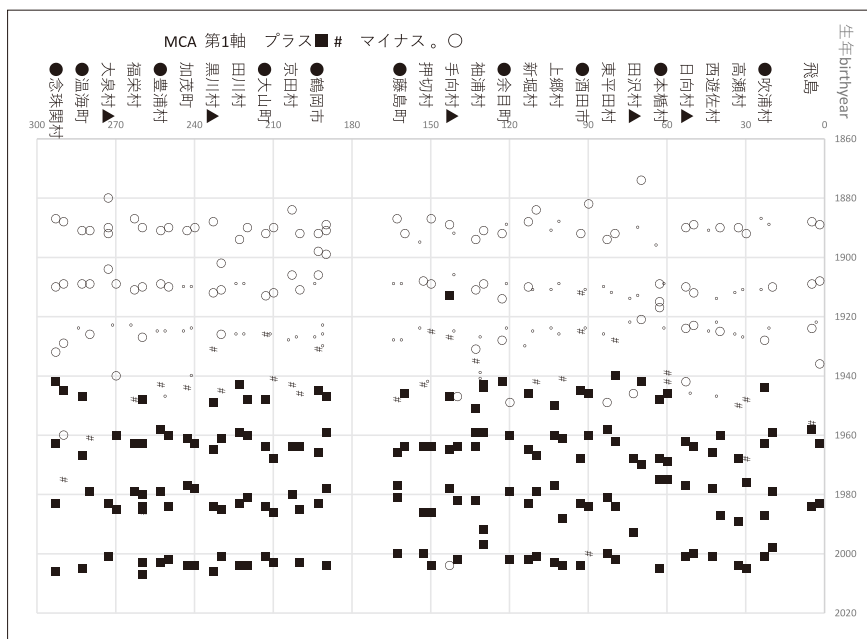


図2 サンプルスコアのグロットグラム 第1軸

化があったと思われる。140年の年齢差，南北80kmほどの地域差を比べると，多変量解析の結果の第1軸として，年齢差が析出されたことになる。浜荻語形が若い人に多いのは奇妙だが，「ちっこい」のように，例外的に東京俗語と一致する語形が増えたパターンである。他の図で確認すると，東京の口語形「ゆうがた・あした・ひるま・あったまる」などがある。戦前の標準語教育でも戦後の共通語教育でも，教科書や文章語の格式の高い語形が普及した。20世紀末期からはテレビや実際の人の行き来を通じて，東京の口語や俗語が庄内に入った。それが第1軸に反映したのだろう。

2.2 多重対応分析の個人ごとの値 第2軸

図3に第2軸のサンプルスコアを示す。プラスとマイナスが年齢層によって分けられるが，一方向的ではない。プラスの大きな値の波が1940年前後と2000年前後の2回現れる。また地域差が見られ，南端鼠ヶ関と北端飛島に目立ち，鶴岡付近にプラス値は少ない。地方的周囲分布の古形残存に対応する。250年前の浜荻語形が，辺境に残っていたが，第2次調査の高年層と若年層の記入で復活したパターンである。第2次調査で調査票の表記が現代かなづかいになり，ふだんのことばとの結びつきがたやすくなった効果だろう。若年層で多くなったのは，21世紀に入り，方言を見直す雰囲気広がったためだろうか。実際に日常使うかの判断基準が中学生と上の世代で違っていた可能性があ

る。(図14の原図で個々の語形を再確認すると，「ねんきあまめ・とんてき・おえさま・ひしゃく・わる」など，現代はほとんど耳にしない語が大きいプラス値をとる。20世紀末期に庄内各地で新たな方言集が多く編纂されたが，そこにも載らないような，廃語・死語である。高年層と中学生が，自分が使わなくとも地域で使う人がいれば「○使う」と記入したと考えられる。)要約すると，第2軸は年齢差と地域差を示す。

2.3 多重対応分析の個人ごとの値 第3軸

図4に第3軸のサンプルスコアを示す。プラスが年齢軸の真ん中に割り込む不思議なパターンである。北部ではほぼ全年齢層がプラスだが，南部では第1次調査の高年層に多い。250年前の浜荻語形が，第2次調査では，北半高年層に残っていた。第2次調査では全体の高年層で「使う」が増えた。調査票の表記が現代かなづかいになった効果だろう。若年層では減った。

常識的な共通語化，方言の衰滅を示す。

要約すると，第3軸は年齢差と地域差を示す。(図15の原図で語形を再確認すると，「やみさお・やみ・ななばち・てのこっぱ・じくねる」など，現代はほとんど耳にしない語で，庄内北半に分布する語が大きいプラス値をとる。逆にマイナス側には「だんま・かいかいろう・くしびきずみ」のように鶴岡周辺部に分布する語がある。)

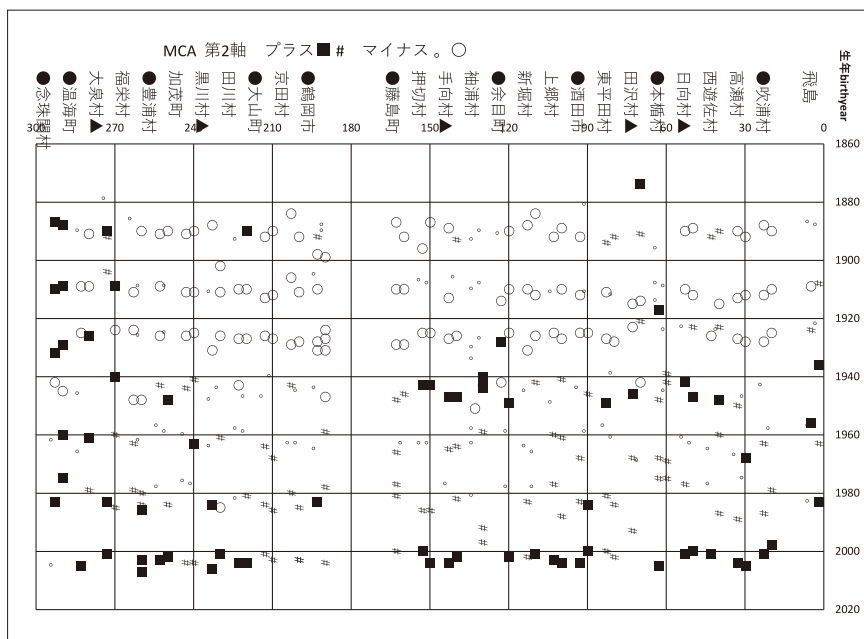


図3 サンプルスコアのグロットグラム 第2軸

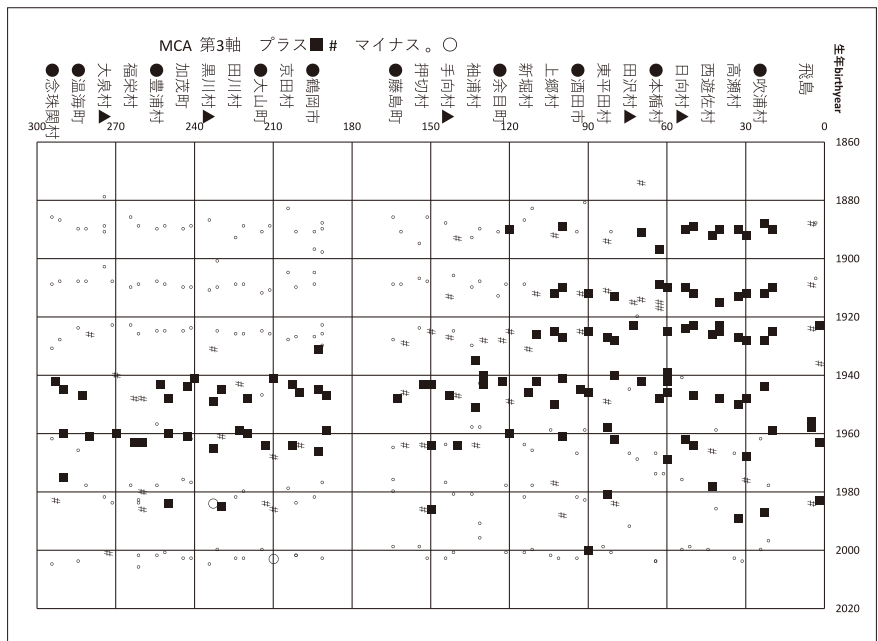


図4 サンプルスコアのグロットグラム 第3軸

2.4 多重対応分析の個人ごとの値 第4軸

図5に第4軸のサンプルスコアを示す。この軸も年齢差と地域差を示す。第3軸と似て、南北の地域差がある。世代差は、第2次調査の高年層で一時値が低くなるという不思議なパターンを示す。第3軸同様第1次調査では北部のみで使われた語形が、第2次調査では全域の若い世代に広がった。図15で見るように、3

軸プラスの語は右上にあり、4軸プラスの語と重なる。まれにしか使われない言い方が、特別なパターンとして現れたのだろう。

以上と逆なのが、第3, 4軸マイナスの語である。具体例は「だんま」で、南部の鶴岡付近に多い語で、当時から分布は範囲を広げていないか、中心都市から今でも進出中の語である。

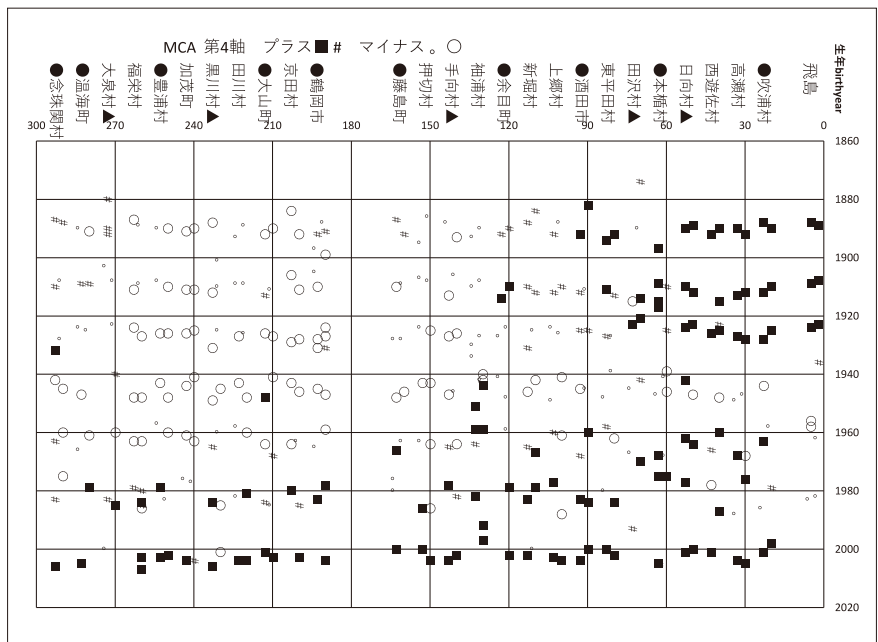


図5 サンプルスコアのグロットグラム 第4軸

まとめると、第3、4軸は第1、2軸で説明できなかった残差を扱うもので、小さな分布傾向を取りだすものだった。第2軸で南北差と周囲分布が分離されたが、その傾向をさらに細分するものだった。

以上個人ごとの値を図化することによって、庄内地方の語彙変化の主な様相を、大きく把握できた。以下では、同じサンプルデータの平均値を計算して、大傾向を把握しよう。

3. インフォーマントの地域差 分布地図

以上の単純化グロットグラムによるMCA値の表示では、個人ごとの値を図化して考察できた。以下では4軸それぞれのサンプルスコアの地点別平均値を地図上に示す。図6～9では10段階以上に細分して記号の大きさで示した。なお図2～5では値を4段階にまとめて示した。

3.1 地点ごとの値 第1軸 年齢差と酒田

図6にサンプルスコア第1軸の値を地図化した。地理的な差があり、港町で商工業都市の酒田とその付近のプラス値が大きく、周辺部で小さい。酒田中心の分布である。東京の俗語が、港町酒田付近から普及しつつある。最小値は飛鳥で、城下町鶴岡の値も小さく、地図の■印は見落とすほどである。(図2で個人ごとに見ると、年齢差が激しく、戦前生まれはマイナス値だったが、地点ごとに平均値を出すとプラスになる。)

3.2 地点ごとの値 第2軸 鼠ヶ関の古形残存

図7にサンプルスコア第2軸の値を地図化した。(図7～9は半沢(2021)にも提示した。)庄内地方南西端鼠ヶ関付近でプラス値が大きく、その東の山間部の

村もプラス値を示す。他で小さく、マイナス値は鶴岡付近と北部に分布する。庄内南端と他の対立である。(図3に戻って個人ごとに見ると、年齢差が大きい、上の世代と下の世代の値が釣り合って、平均値としては、鼠ヶ関付近以外の違いが見えない。)グロットグラムを作って個人ごとの値を表示したのは、半年以上経ってだが、図7を使った初期の分析からは、納得の行く結論が得られなかった。この軸では地域差が出ないので、地図化する必要がなく、グロットグラムで十分である。むしろ図3のようなグロットグラムで地理と年齢を同時に観察するほうが実りが大きい。鼠ヶ関付近では、のちに東京の俗語として庄内に広がる語形を、近代まで保っていたと解釈できる。

406枚の年齢柱地図を個別に通覧したときには、鼠ヶ関の特殊性に気づかなかった。多重対応分析の第2軸で取り出されたことで、意義に気づいた。鼠ヶ関には古代以来海沿いの関所があった。江戸時代の庄内藩などの公的な地図では、鶴岡と新潟県北部の村上との間の道は海沿い(鼠ヶ関を通る現在の国道7号線)とその東の山中(現在の国道345号線)の2経路が記入してある。しかし江戸時代の庶民向けの道中図などでは、東の山間部の峠を何度も越える道だけが書いてある。羽越線開通(1920年代)以前の徒歩の旅では、鼠ヶ関は鶴岡から遠い集落だった。海沿いの道は平坦だが、岬を回ると遠いし、嵐のときは波風が不快で、危険でもある。鼠ヶ関の方言残存率の高さは、鉄道未開通時期の交通の不便さの反映として説明できる。図10～13で見ると、鶴岡から鼠ヶ関への徒歩距離は、庄内平野北端の日向(にっこう)村、本楯村に匹敵するほど遠い。羽越線が多数のトンネルを掘りながら海沿いルートを取ったのは、冬の豪雪が海沿いにはないことが理由という。鉄道開通以来庄内平野の交通は変わり、

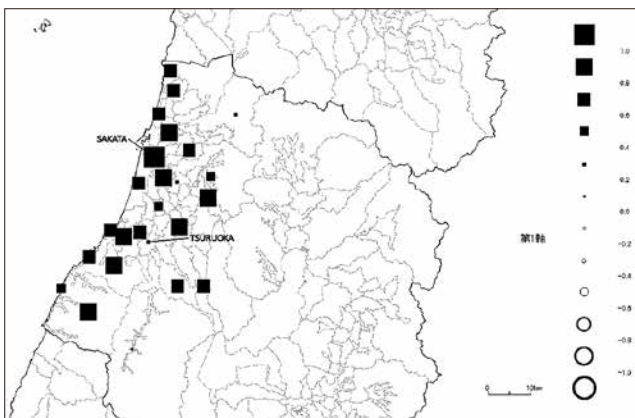


図6 サンプルスコアの分布地図 第1軸

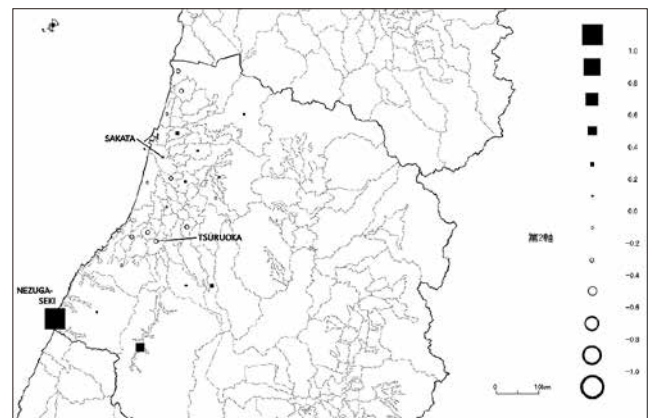


図7 サンプルスコアの分布地図 第2軸

メンタルマップも鉄道中心になった。これが方言分布に反映したわけである。

3.3 地点ごとの値 第3軸 南北差

図8にサンプルスコア第3軸の値を地図化した。地域差が顕著で、飛島を含む北部にプラス値が大きく、鶴岡付近はマイナス値が多い。北端吹浦がプラス値最大。マイナス値が南半分の鶴岡付近に多く、大山が最小である。単純な南北差ではなく、鶴岡付近と、徒歩距離の遠い北端との対立と読み取れる。(図4に戻ると、第1次調査で北半でのみ保たれていた浜荻語形が、第2次調査時には南半でも認められるようになった。)復活したのだろうか。表記が現代かなづかいになったために、南半の高年層も使用を認めるようになった可能性がある。「だんま」の具体例で確認できる。前述のように、第3,4軸のプラスの語形では、忘却寸前、消滅しかけの危篤状態の語が、極端なプラス値を示す。

3.4 地点ごとの値 第4軸 南北差

図9にサンプルスコア第4軸の値を地図化した。こ

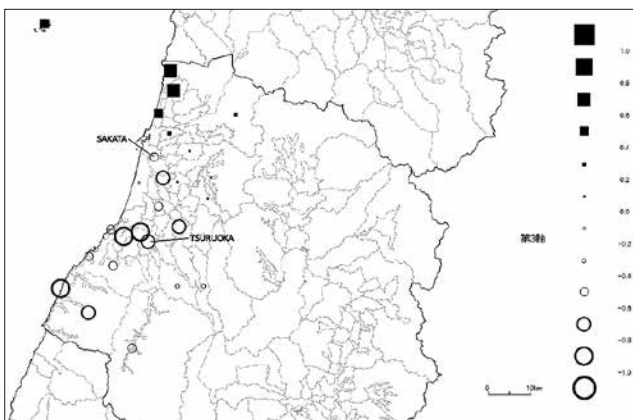


図8 サンプルスコアの分布地図 第3軸

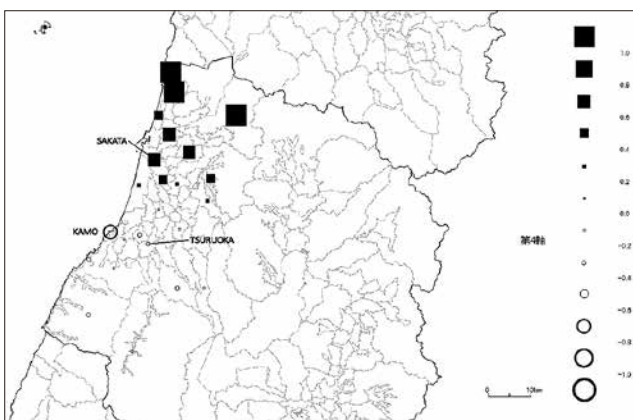


図9 サンプルスコアの分布地図 第4軸

れも地域差が顕著で、再び南北の対立だが、北の吹浦付近でプラス値が大きく、日向村が最大。南部がマイナスで、港町加茂が最小である。北半分に浜荻語形が残ったと読み取れる。(図5に戻ると、第1次調査で北半でのみ保たれていた浜荻語形が、第2次調査時に酒田付近の中年層をピークに若い世代に普及している。250年前の浜荻語形が若い世代に再進出している。年齢層の区切りは調査票の表記変更と関係が薄そうに見える。図15の原図によって具体的な語形を見ると、第3,4軸のプラスの語形もマイナスの語形も、大きな値を取るものは、現在では耳にしない古めかしい方言である。)方言の復権の波に乗って再び活力を得たのか、さらなる考察が必要である。

以上のように、多重対応分析の結果は、インフォーマントの値の平均値を年齢差で見ても、地域差で見ても、論理的に説明でき、かつ個々の単語の分布図(年齢柱地図)と照合しても、対応を示すものだった。個々の分布図を通覧しても分からず、総合図で見ても気づかないようなパターンを発見できた。全406項目の内部構造を見たことになる。

4. 徒歩距離と多重対応分析の4軸の値 散布図

以下では考察の次元数を単純化するために、鶴岡からの江戸時代の道路での徒歩距離による地表上の位置を図化する。散布図の技法を用い、横軸に鶴岡からの徒歩距離をとり、縦軸にMCAのサンプルデータの地点ごとの(男女7世代の)平均値を連続的数値として示す。地理的位置は東西については捨象した。

以下では情報を補うために●▲で地表上の位置を示した。4枚の図を比較しやすくするために、軸の最大値、最小値を2.0から-1.0に統一した。井上・半沢(2021.1)では、第1次と第2次のデータを分けてMCAを適用した結果について述べたが、この稿では全体に適用した結果について述べる。

4.1 多重対応分析と徒歩距離 散布図 第1軸

図10に第1軸の値による散布図を掲げる。最大値を示すのは酒田市で、南端の村々と飛島は小さい値である。近似曲線は、中央部で高く周辺で低い。酒田付近で多く使われる浜荻語形を示す。つまり今でも庄内全体に普及中の語形である。冒頭に●のついた地名は、羽越線の駅のある市町村で、近似曲線の上に位置することが多い(鶴岡・余目を除く)。末尾に▲のついた

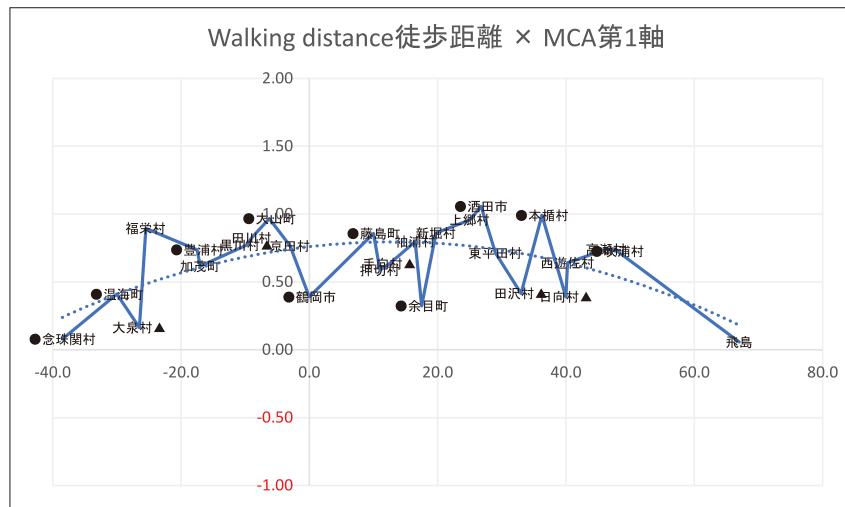


図10 MCAと徒歩距離 散布図 第1軸

地名は庄内平野東部の山間部に近い村を示すが、ほぼすべて近似曲線の下である。つまり浜荻語形をあまり使わない。第1軸プラスのグループは、年齢差が大きい語だった。若い世代で増えた東京の俗語で、1940年代生まれ以降に（再）普及した。港町、商業都市酒田が先頭を切って、鉄道沿いに早く伝わり、周辺の山村には普及が遅れている。

4.2 多重対応分析と徒歩距離 散布図 第2軸

図11に第2軸の値による散布図を掲げる。前図と逆に中央部の値が低く、周辺部で高い。(図3の個人値に戻ると、第2軸プラスは南端と北部に多い。第2次調査の1940年代生まれと若い世代に多いのは、方言復権の動き、または中学生たちが積極的に使うに印をつけたためかとも思われる。) 鶴岡付近から浜荻語形が

駆逐され、周囲分布を示し、徒歩距離の遠い場所に残存している。これが地点ごとの平均値でも、読み取れた。

4.3 多重対応分析と徒歩距離 散布図 第3軸

図12に第3軸の値による散布図を掲げる。前図と違って北半の値が大きい。(図4によれば、確かに北部に多い。これが地点ごとの平均値でも、読み取れた。)

4.4 多重対応分析と徒歩距離 散布図 第4軸

図13に第4軸の値による散布図を掲げる。前図と同様北半の値が大きい。地点による平均値の違いが大きい。(図5によれば、確かに北部に多い。かつ若い世代に多い。これが地点ごとの平均値でも、読み取れた。)

第1～4軸の地点ごとの平均値をまとめると、徒歩

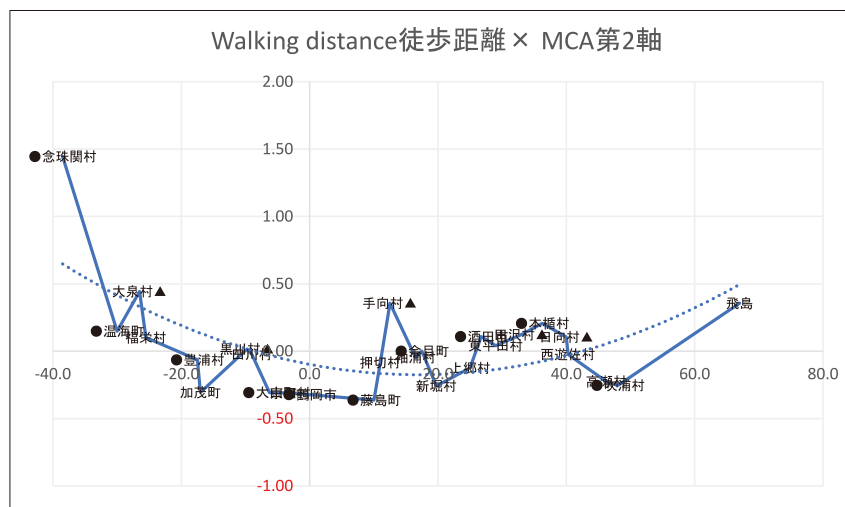


図11 MCAと徒歩距離 散布図 第2軸

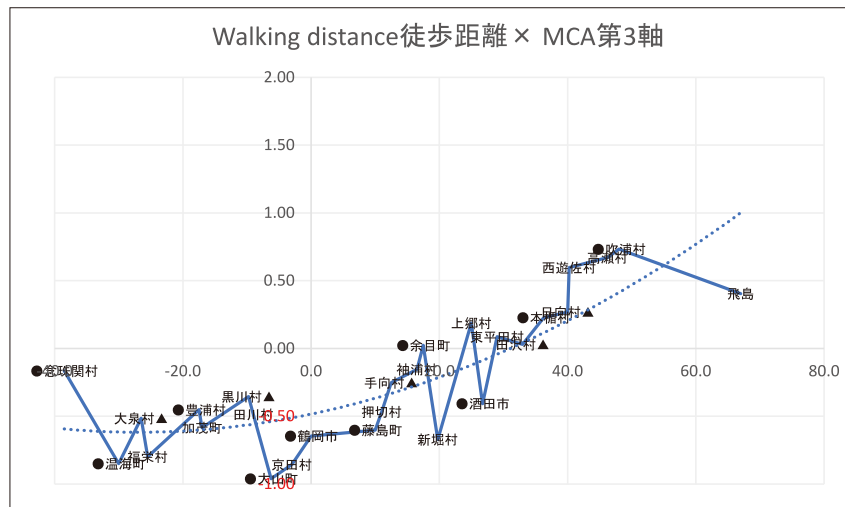


図12 MCAと徒歩距離 散布図 第3軸

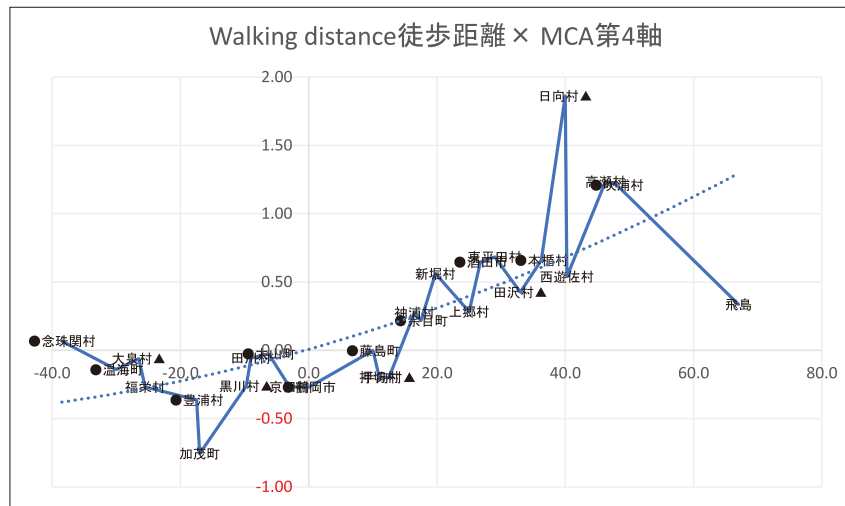


図13 MCAと徒歩距離 散布図 第4軸

距離は、4軸すべてと何らかの相関を示す。単に南北だけでなく、東西つまり東の山間部との関係を見ると、第1軸では影響が見られた。さらに近代の鉄道(昭和初期1920年代の羽越線開通)によって、沿線のことばが変化したことが読み取れた(井上・半沢2021.9)。ことばの変化に人々の交流、対面の会話が作用することを物語る。鶴岡からの江戸時代のルートによる道路を手がかりにしたが、現在の方言分布(浜荻語形の残存度)を説明できることが分かった。

5. 浜荻語彙のMCA分析結果

5.1 MCA 第1軸×第2軸

以上で述べたのは、データ行列のうち、人に関わる次元の分析だった。以下では語に関わる次元の分析に

ついて論じる。ここにはInoue & Hanzawa (2021.9)と同じ図を再掲する。地理的分布パターンの典型8項目の語形をラベルで示し、他の項目はラベルなしで記号のみをプロットした。なお井上・半沢(採用決定)には項目名を添えたグラフを掲げた。

図14によれば、横方向第1軸のプラス側には「ちっこい・やっぱし」などの、今も若い世代に盛んに使われる語形が集中する。マイナス側には「わにる・ほちゃ・やみ・けとぎ」などの、高齢者にかろうじて残る語形がある。第1軸は年齢差を反映し、これまでの人の行の分析と一致する。縦方向第2軸のプラス側では「わにる・ほちゃ」が大きな値を取る。北部に分布する語形である。「けとぎ」がマイナスで、その項目の年齢柱地図とグロットグラムによれば(Inoue & Hanzawa 2021.9)、第1次調査と第2次調査の違いが

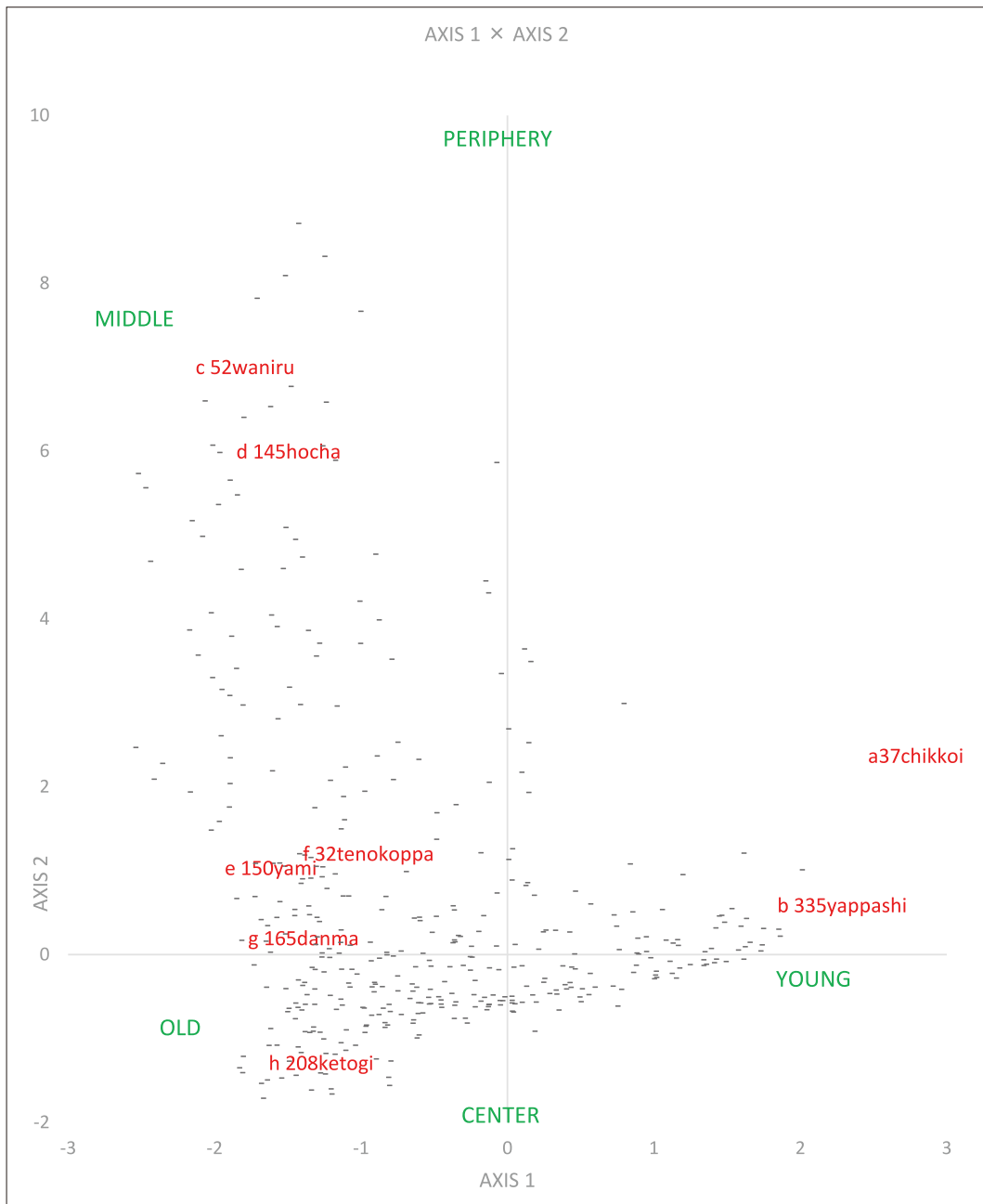


図14 MCA 第1軸×第2軸

激しく、かつ鶴岡付近に多く、山間部には少ない。実物の普及時期との関係が考えられる。第2軸は鶴岡などの平野中央部と周辺部とを分けると解釈される。

5.2 MCA 第3軸×第4軸

図15によれば、横方向第3軸と縦方向第4軸は、ほぼ同じ傾向を析出している。第3軸と第4軸ともにプラス側の第1象限には「やみ」と「てのこっば」がプロットされた。庄内北部に分布する古形である。第3軸と第4軸ともにマイナス側の第3象限には「だんま」

がプロットされた。鶴岡周辺に分布する。第3軸と第4軸は、鶴岡周辺と庄内北部を分ける働きをする。上記、人の行で得られた結果と同じである。

以上、MCAの結果の単語の列を考察して、(当然ながら)人の行と対応することを確認した。図14と図15では典型的8語のみをラベルで示したが、付近にプロットされた他の語形は似た地理的分布を示す。また原点付近の語形は、特徴的な分布パターンを呈しない。

多重対応分析の適用により、400語以上の多数の語全体の分布を大きく把握することができた。「語はそ

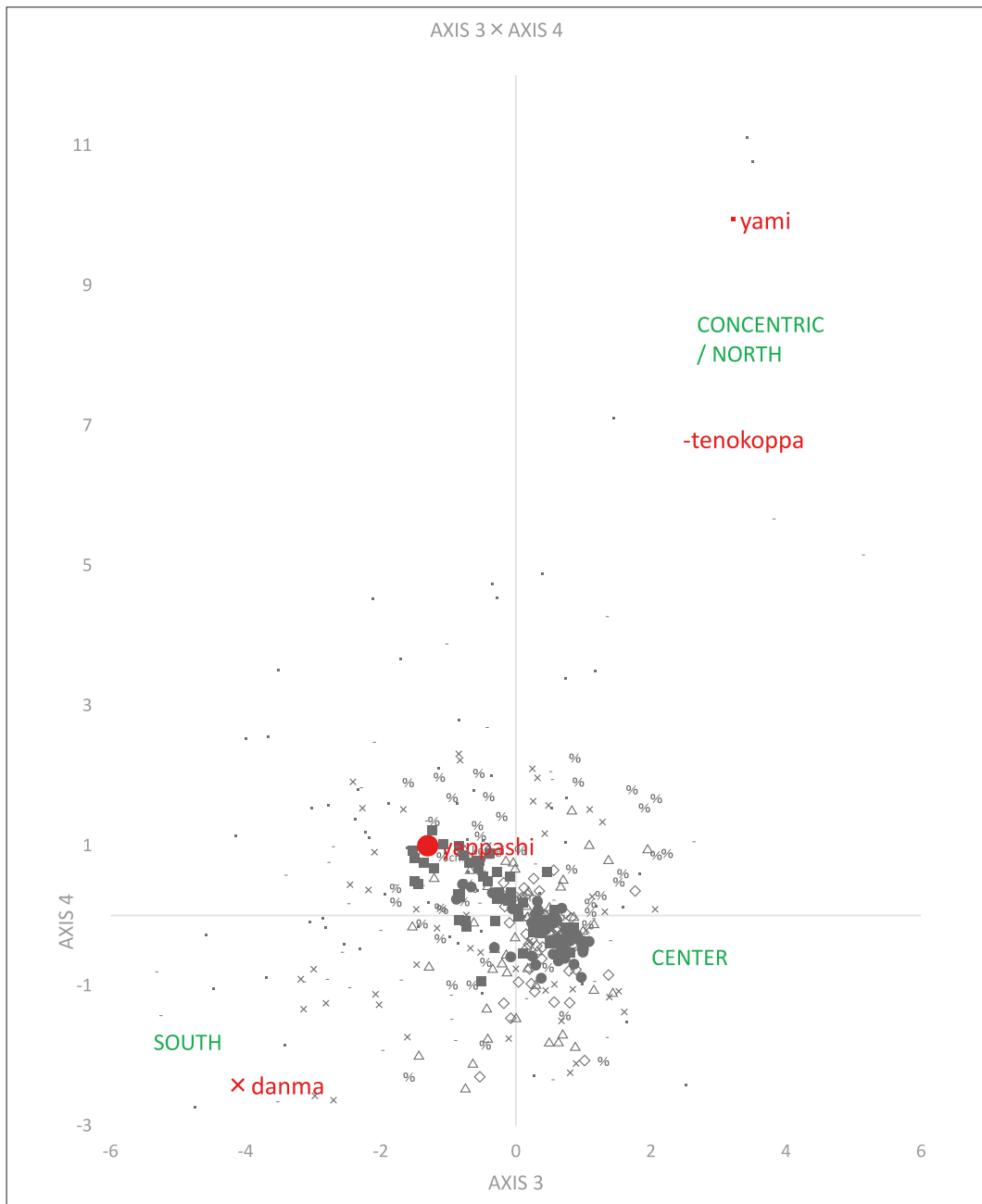


図15 MCA 第3軸×第4軸

れぞれ独自の歴史を持つ」(Jaberg 1908) という言が方言学では有名だが、多数の語を計量的に扱くと、一定の傾向が浮かびあがる。浜荻語形でも成功した。

6. 浜荻語彙の時間と距離

以上の結果を別の目で見直そう。浜荻の年齢柱地図では、1950年の3世代調査を活用できたので、合わせて140年の世代差が読み取れた。地理的な広がりも調査できたので、3Dのグロットグラムと解釈できる。

これに多変量解析を施した結果は、これまでの線状の地域の年齢差70年程度のグロットグラムとは別の価値を有する。

6.1 年齢差が大きい

大きな成果は年齢差が大きいと分かったことである。140年という長さが作用し、第1軸で表れた。調査全体としては、年齢差が大きいのが、これは庄内地方という規模のためだろう。日本全体でもLAJのインフォーマント(19世紀末生まれ)から今生まれて

いる子どもまでは100年以上の違いがある。浜荻語形は、失われつつある。ただしほぼ全員が今も使う語もある。現代口語として使われる語である。共通語化は完成直前と言ってよい。しかし共通語化でない言語変化が、空からばらまいたような分布でなく、隣接地域から地伝いに伝播した場合、ことに新方言の場合は、年速1km前後で伝播したら、新たな地域差が生まれる。新方言の普及過程は今でも観察可能であり、浜荻調査でも、いくつかの項目で見られた。多変量解析による総合的分析を踏まえて、さらに解析を進める必要がある。

1940年代生まれの人に急速な共通語化があったが、この時期は鶴岡市共通語化調査の音韻4回の調査結果と似る。しかしアクセントの共通語化は遅れた。文法の共通語化も遅れ、むしろ新方言の普及が観察される。また音韻についても、カタ行子音の語中での有声化は強固に残っている。年齢差に基づく変化時期を見るには、庄内を離れて、広い地域での考察が必要である。

テレビの大規模普及は1962年の東京オリンピックの前後であり、1940年代生まれの世代は、言語形成期以降に接したことになる。テレビ漬け世代より前に語彙変化が進行していたことになる。

6.2 第1次調査の地域差と第2次調査の年齢差

井上・半沢(2021.1)では、第1次調査と第2次調査に分けて、多重対応分析を適用した。第1次調査では第1,2軸が南北の、第3軸がその下位の地域差を示した。第4軸で初めて年齢差が登場する。第2次調査では第1,2軸は年齢差が大きく、第3,4軸は辺境と中央という地域差が大きい。つまり2回目の調査では地域差が薄れて、年齢差が大きくなった。これは、Inoue and Hanzawa(2021.9)に掲げた個々の語の分布図の、7年齢層男女の円グラフでも分かる。若い世代が急速に浜荻語形を使わなくなっている。

6.3 変化・伝播の連続体

個々の項目の地理的年齢の様相は、連続体を示す。「じじ」と「ばば」のように、非常によく似た分布を示す図もある。しかし多くの場合微妙な違いがある。方言地図で千差万別の地理的分布が見られたと同様に、グロットグラムでも多様な分布が見られる。見方を変えると、多くの図は全体として、変化・伝播の総合的過程を示すと、とらえられる。ランダム過程を含みながら、項目により、人により少しずつずれながら変化が進む。

項目ごとに言語的にみれば、意味分野、実物の消滅年代、使用頻度、使用場面などが影響する(井上・半沢2021.4)。地点・個人ごとに非言語的にみれば、中心都市との交流、交通路、産業構造、住民の移動、などの要因が左右する。本稿では中心都市からの徒歩距離・所要時間、鉄道路線沿いか、山間部かなどの要因にも配慮した。もっと多くの要因を考慮に入れられてしかるべきである。さらに同一地点の住民であっても、年齢以外に性別、職業、地理的・社会的活動状況(ネットワーク)が働く。さらに個人的要因として、応化accommodationの大きさ、ことばについての信条beliefも影響する。

ただし今回のように地理的・年齢的変異の大きい集団では、それ以外の要因の働きを抽出するのが難しい。疑いなく証明するためには、他の条件を一定に統制した上での客観的分析が必要である。関連する研究を参照する必要がある。

6.4 方言の全体像

計量方言学の発達によって、方言分布の全体像把握が容易になった。様々な多変量解析が適用されたが、その多くは地点間の類似度を扱うもので(Heeringa 2004, Goebel 2020)、項目間の言語的情報は考察されない。代わりに直線距離・道路距離・過去と現在の移動時間などの非言語的要因との相関が計算・検討された(Jeszenszky et al. 2019)。その点日本で適用されてきた林の数量化第3類とその後欧米で普及した対応分析は、地点の次元と言語項目の次元の双方を同時に分析できる点が有利である。初期に比べてコンピューターの性能が大幅に向上し、多数の項目を一度に分析できるようになった(鎌水2017)。ここでも江戸時代の方言集収録語全体の400語以上を分析でき、大きな傾向を把握できた。今後日本全体について、多数の項目を元にした分析が進み、方言の全体像が把握されることが期待される。

6.5 方言の将来

グロットグラムでは現代若年層の方言分布が分かる。これまでの全国高校調査でも若年層の状況が分かる(井上2000)。庄内4高校調査でも(井上2000)、鶴岡・酒田の新しい方言差が生まれている。内陸地方、山形市の影響もある。新方言が活躍していて、メンコイ(かわい)では、『浜荻』成立以来250年に数段階の変化が見られた(Inoue 2020.12)。また方言と地理、社会との密接な関係が読み取れた。鶴岡市では20年間隔の

共通語化調査が4回行われており、将来もことばの定点観測が企画されている。この『浜萩』グロットグラムにより、過去に延長し、成立以来250年の歴史と結びつけることが可能になった。

7. 結論 年齢柱地図からグロットグラムへ

庄内浜萩方言地図は、2次元の地表面に年齢の柱を立てた3次元のグロットグラムとして構想された。406項目、372名全体の多変量解析を通して、第1軸は年齢差を強く反映し、第2, 3, 4軸は主に地域差を反映すると、解釈できた。つまり140年の年齢差は庄内地方の南北約80kmの地域差より大きい。なお東海道グロットグラム調査(60年の年齢差と約600kmの鉄道距離)の林3類分析結果では、第1, 2軸とも地域差を示した(井上2001)。庄内方言の地域差の基本は、第2, 3, 4軸ともに南北の対立と見られた。大まかには南北の次元だけでも近似できる。3次元の「年齢柱地図」として企画した研究の結果DASH Dialect Atlas of Shonai Hamaogiは、東西の隔たりを捨象し、従来どおりの線状の地域の2次元の「単純化グロットグラム」としても大まかには近似できる。次元数が少ないだけに、年齢差と地域差を効率的に把握できる。全国方言の地表上の分布が、東京または京都からの鉄道距離を使えば、線で近似できるのと、同様である(井上2001)。多変量解析にかけて、データの内部構造を多次元的に分析すると、その結論はもっと単純な図式と算術計算で再現できることがある。今回の研究でも、単純化ができた。

方言地図もグロットグラムも、個々の回答を忠実に(フランス式に文字で、またはドイツ式に記号で)表現することにより、ことばと人間、さらにその奥に潜む心理や地理、社会、経済との関係を研究できた。ただし個別主義atomismと非難されることもあった。一般理論に近づくためには、大量データを総合的に扱って、大きな傾向を把握する必要がある。複雑な現象を単純化することがその手法の一つである。本稿で採用した多変量解析法と、その結果の効率的図示報告としての「単純化グロットグラム」は、その具現化と言っ

【謝辞】

浜萩調査では、現地の教育委員会、コミュニティセンターその他の機関に人探しの労を取っていただいた。また大勢の方々に調査票記入をいただいた。深甚

の謝意を表す。また本研究は下記の科学研究費の補助を受けた。基盤研究(B)16H03420井上史雄「公共用語の地域差に関する社会言語学的総合研究」。

【文献】(Alphabetical)

- Goebel, Hans (2020) Dialectometry
<http://dialektkarten.ch/dmviewer/alf/index.en.html>
- 半沢康 (2017) 「グロットグラム調査データの実時間比較」大西拓一郎編『空間と時間の中の方言』朝倉書店：283-303
- 半沢康 (2018) 「現代における方言語彙の動態」小林隆編『方言の語彙』明治書院：132-147
- 半沢康 (2021) 「『庄内浜萩』調査データの多変量解析」『福島大学人間発達文化学類論集』33：49-64
- Heeringa, Wilbert (2004) Measuring Dialect Pronunciation Differences using Levenshtein Distance. *Groningen Dissertations in Linguistics* 46
- 井上史雄 (2000) 『東北方言の変遷』秋山書店
- 井上史雄 (2001) 『計量的方言区画』明治書院
- 井上史雄 (2018) 「平成の方言——鶴岡の二五〇年間の言語変化——」『日本語学』37 (10)
- 井上史雄 (2019) 「方言語彙変化の140年」『第1回日本地理言語学会予稿集』
- Inoue, Fumio (2020.12) Dialect vocabulary changes over 140 years — Standardization and new dialect forms observed in Hamaogi glossary —. *Dialectologia: et Geolinguistica* 28 : 105-116.
- 井上史雄・半沢康 (2019) 『庄内浜萩方言地図 Dialect Atlas of Shonai Hamaogi (DASH)』科研費報告CD. <http://www.urayasu.meikai.ac.jp/japanese/inoue/inouetop.htm>
- 井上史雄・半沢康 (2021.1) 「方言語彙の地域差と年齢差——庄内浜萩調査2回の多重対応分析——」『福島大学地域創造』32(2)：5-24
- 井上史雄・半沢康 (2021.4) 「方言衰退の語彙論的過程——庄内浜萩の250年——」『日本語の研究』17 (1)：1-18
- 井上史雄・半沢康 (2021.9) 「方言語彙残存と鉄道交通——『庄内浜萩』の方言語彙残存率——」『社会言語科学』24(1)：144-156
- Inoue, Fumio and Hanzawa Yasushi (2021.9) Multivariate analysis of geography and age in dialect vocabulary — Comprehensive analysis of 250 years of language change —. *Dialectologia:*

revista electrònica 20 : 97-160

井上史雄・半沢康（採用決定）「方言の地域差から年齢差へ——庄内浜荻追跡調査の多重対応分析——」『言語研究』

井上史雄・半沢康（未刊）DASH GL.

Jaberg, Karl (1908) *Sprachgeographie*. Aarau (p.6 p.22).

Jeszszsky, Péter, Yoshinobu Hikosaka, Satoshi Imamura and Keiji Yano (2019) Japanese Lexical Variation Explained by Spatial Contact Patterns. *International Journal of Geo-Information* 8 (9), 400 ; <https://doi.org/10.3390/ijgi8090400>

洪惟仁（2019）『台湾社会語言地理学研究 I II』前衛出版社

国立国語研究所（1953）『地域社会の言語生活——鶴岡における実態調査——』秀英出版

Mang, Alexander and Markus Wollin (2010) Sprachraum and sociodemographic variables. in Lameri et al. (2010) *Language and Space*, 607-626, Maps 3101-3106. Berlin: Mouton de Gruyter.

三矢重松（1930）『庄内語及語釈』刀江書院.

Szmrecsanyi, Benedikt (2008) Corpus-based dialectometry: aggregate morphosyntactic variability in British English dialects. *International Journal of Humanities and Arts* bszm@frias.uni-freiburg.de

鎌水兼貴（2017）「現代日本語の共通語化過程」大西拓一郎編『空間と時間の中の方言』朝倉書店

in the scatter plot. In addition, the difference in vocabulary by age and locality for each of the 370 informants was displayed. As a result, the difference according to age was most significant. The greatness in length of 140 years appeared on the first axis. The second, third and fourth axes all pointed to the basis of the regional differences between north and south. When many words are processed computationally, theoretically useful tendencies can be grasped.

Keywords Glottogram, Multiple Correspondence Analysis, Standardization, Hamaogi

【注】

1 『浜荻』所収の俚言406語。調査・分析の対象。

〔原稿受付(2021年12月13日), 原稿受理(2022年1月7日)〕

【英文要旨】(Abstracts)

“Hamaogi” is a dialect glossary of Tsuruoka in the Shonai district of Yamagata Prefecture, which was compiled in 1767. In 1950 and 2018, the Hamaogi Glottogram Survey was conducted to examine the residual rate of vocabulary. The data consists of seven age groups at intervals of about 20 years, so the maximum difference between the youngest and eldest groups was about 140 years. About 370 people from 27 localities shown in 27 pillars distinguished by age difference were plotted in “age pillar maps”.

Multiple Correspondence Analysis was applied to 406 word forms and the values obtained for those forms on the first to fourth axes were shown