

国立国語研究所学術情報リポジトリ

話者分類に基づく地域類型化の試み： 全国方言意識調査データを用いた潜在クラス分析による検討

メタデータ	言語: Japanese 出版者: 公開日: 2015-10-30 キーワード (Ja): キーワード (En): common language, regional dialect, language consciousness, regional typology, latent class analysis 作成者: 田中, ゆかり, 前田, 忠彦, TANAKA, Yukari, MAEDA, Tadahiko メールアドレス: 所属:
URL	https://doi.org/10.15084/00000493

話者分類に基づく地域類型化の試み

——全国方言意識調査データを用いた潜在クラス分析による検討——

田中ゆかり^a 前田 忠彦^b

^a日本大学／国立国語研究所 共同研究員

^b統計数理研究所／国立国語研究所 共同研究員

要旨

方言使用にかんする地域類型を考察するために、16歳以上の男女を対象とした全国規模の言語意識調査のデータを分析した。我々は言語使用にかんする話者個人レベルでの確率的なクラスタリングを得るために、潜在クラス分析を利用した。その結果に基づいて、地域的分布の特徴を調べることで得られたクラスター（潜在クラス）と地域との対応関係を同定することを試み、またその他の対象者属性と潜在クラスとの対応関係を精査することにした。その結果、次の五つの潜在クラスが抽出された：「クラス1：積極的方言話者」「クラス2：共通語話者」「クラス3：消極的使い分け派」「クラス4：積極的使い分け派」「クラス5：判断逡巡派」。用いた説明変数のうち、クラス帰属への効果が有意となったものは効果の大きな順に、生育地、職業、教育程度、年代であった。居住地都市規模と性の効果は有意ではなかった。話者の生育地の観点から、各クラスの特徴を示すと次の通りとなる。「クラス1：近畿・中国・四国生育者」「クラス2：首都圏・北海道生育者」「クラス3：北関東・甲信越・北陸・東海生育者」「クラス4：沖縄・九州・東北・中国生育者」「クラス5：北海道生育者」。職業についてもいくつかのクラスで特徴的なパターンを示した。年齢効果は、全体のサンプルでも、またいくつかの地域別に分析した結果でも、非線形な関係を示した。概してこれらの年齢の効果は二つかそれ以上の変化点をもち、多くの場合に35–40歳前後と60歳周辺に観察される。以上のような結果に基づき、我々の分析で得た地域類型の実質上の意味や、先行研究で得られた類型との関係を議論した*。

キーワード：共通語、方言、言語意識、地域類型、潜在クラス分析

1. はじめに

1.1 話者の分類に基づく地域類型化を試みる背景

日本語社会における「方言」や「共通語」に対する意識は、時代によって大きく変化してきた（田中 2011: 40–66）。時代をこんにちに限定して考えてみても、話者の生育地や居住地によって、その意識は大きく異なることが予想される。たとえば、生育地や居住地が「方言主流社会」なのか、「共通語中心社会」なのか（佐藤 1999）。さらに、その「地域」や「地域で使用されている方言」が日本語社会においてどの程度の「威光 (prestige)」をもっているのか。加えて、話者の年代や性、職業や教育程度、ことばに対する志向性などによっても異なるだろう。

このような「方言」や「共通語」に対する意識の違いは、個人個人の言語行動に対するスタンスやその言語実態にも大きな影響を与える。それと同時に、「方言」や「共通語」に対する意識は、一定の広がりをもつ「地域」によって異なるパターンが存在し、しかも、そのパターンはある程

* 本稿は、国立国語研究所基幹型共同研究「多角的アプローチによる現代日本語の動態の解明」（プロジェクトリーダー：相澤正夫）の一環として実施した調査研究に基づくものである。

度類型化できそうである。

実際、「方言」や「共通語」に対する言語意識や、それらにかんする言語行動パターンから、地域や話者を分類し、類型化しようとする試みは、様々な角度から試みられてきた。研究者の観察に基づく類型化にはじまり、時代や地域、調査対象の異なる大規模な意識調査データに基づく類型化まで、いくつか存在している（1.2節参照）。

本稿では、2010年に実施した大規模な無作為抽出データに基づく「方言」と「共通語」にかんする意識調査データを用いて、話者（回答者）を分類し、そこから地域類型の抽出を試みる。話者ひとりひとりの分類というミクロな視点から地域類型を捉えようというものである。話者の分類と地域類型の抽出には、多変量解析の手法の一つである潜在クラス分析を用いて、研究者の主觀に依らない地域類型のこんにち版を提案することを目的とする。

その上で、本稿と同じデータを用いて地域ごとの回答パターンに基づく地域の類型化を目指した田中（2011: 92–114）を中心に、「方言」と「共通語」に対する意識に基づく地域や話者の類型にかんする先行研究との比較を行なう。

1.2 言語意識からみた地域や話者の類型にかんする先行研究

「方言」と「共通語」にかんする言語意識から、地域や話者を類型化しようとしたものは、研究者の内省や観察に基づく質的データから考察されたもの、質問紙に基づく調査データから帰納的に導かれたものなどがある。ここでは、先行研究が示した類型について概観する。

1.2.1 研究者の内省・観察に基づく類型

研究者の内省・観察・質的データの分類に基づく類型として寿岳（1978）、野林（1971）、陣内（1996: 101–105）がある。寿岳（1978）は内省と観察に基づくもの。野林（1971）は、少年・少女の移住に伴う言語使用意識についての観察に基づく話者の類型を示している。陣内（1996: 101–105）は、九州の大学に通う大学生（九州・中国・関西地域の生育者）の「私の言語生活」「私にとって方言とは」というレポート200通に記述された意見を分類し、方言使用にかんする類型化を試みたもの。以下に、類型①～③として示す。〔 〕内の説明は筆者による要約。

【類型①（寿岳 1978）】

- 「関西型〔地元でもよそでも関西弁〕」
- 「東北型〔地元では東北弁、よそでは共通語〕」
- 「沖縄型〔地元でもよそでも共通語〕」

【類型②（野林 1971）】

- 「支配型〔「規範ことば」「学校ことば」「家ことば」いずれも移住元の方言〕：「首都圏的文化型の地方での実現」
- 「屈服型〔「規範ことば」「学校ことば」いずれも移住先の方言、「家ことば」は移住元の方言〕：「東

日本的な文化型の首都圏での実現」

「反抗型〔「規範ことば」は共通語、「学校ことば」は移住先の方言、「家ことば」は移住元の方言〕」:
「西日本的な文化型の東日本での実現」

【類型③（陣内 1996）】

「無意識顕在化型〔移住先でも生育地方言の使用を継続し、生育地がどこか隠そうとしないタイプ〕」: 関西（特に大阪）、北九州市、福岡市

「意識的潜在化型〔生育地方言を出すまいと意識し、生育地を隠そうとするタイプ〕」: 南九州（鹿児島・宮崎）、出雲

「（「無意識顕在化型」と「意識的潜在化型」の中間型」: 準関西系、準博多系、準北九州系、広島、岡山など

1.2.2 質問紙調査による大規模データに基づく類型

1990年代以降になると、質問紙調査を用いた大規模データに基づく帰納的な話者や地域分類を目指したものが現われてくる。陣内（1999: 135–141）、NHK放送文化研究所編（1997）、田中（2011: 92–114）が該当する。

陣内（1999: 135–141）は、全国14都市に居住する高校生から70代までの約2,800人を対象とした「方言」と「共通語」についての意識調査データに基づいたもの。「地元の道端で同郷の知人と話をする」「東京の電車の中で同郷の知人と話をする」「地元の道端で共通語を話す見知らぬ人と話をする」「東京で共通語を話す見知らぬ人に道を尋ねる」という四つの場面で「共通語」を使うか「方言」を使うかを尋ね、場面による都市別の「共通語」と「方言」使用率の平均値の偏差から導き出された分類である。

NHK放送文化研究所編（1997: 34）には、全国の約30,000人を対象に実施した世論調査「全国県民意識調査」における、「土地のことば」に対する質問項目の回答パターンによって都道府県を分類した表がある。この表は、居住する都道府県の「土地のことば」についての次の三つの質問「この土地のことばが好きだ」（以下、「好き」）、「この土地のことばを残していくたい」（以下、「残す」）、「（この土地の）地方なまりが出るのは恥ずかしい」（以下、「恥ずかしい」）の都道府県別の回答率を全国平均の回答率と比較したもので、47都道府県を全国平均値と比較、分類したものである。ここでは、筆者が、その表に示された「好き」「残す」を「土地のことば」に対する「愛着」と解釈・命名し、整理したものを「全国県民意識調査」による地域分類とみなす¹。

田中（2011）は、2010年12月に層化三段無作為抽出法で抽出した16歳以上の全国の男女を対象に実施し、回答を得た1,341人の全国方言意識調査データを用いたもの。回答者を12の生育地グループに分類し、生育地グループごとの「方言」と「共通語」にかんする七つの質問に対する回答率を変数としたクラスター分析の結果から、大きく二つ細かく六つの分類を示している。

¹ 田中（2011: 95–96）再掲。

以下、類型④～⑥として示す。〔 〕内の説明は筆者による要約。

【類型④（陣内 1999）】

「方言開示型〔どの場面でも方言使用率が相対的に高い都市〕」：京都，東京，札幌，福岡

「方言抑制型〔どの場面でも方言使用率が相対的に低い都市〕」：仙台，千葉，那覇

「使い分け型〔話し相手が同郷人か共通語を話す人であるかによってはっきり使い分けをする都市〕」：弘前，鹿児島，高知，金沢

「中間型〔一貫して平均的な位置にある都市〕」：松本，大垣，広島

【類型⑤（NHK 放送文化研究所編 1997）】

《「土地のことば」に愛着があるグループ》

「好き」「残す」が多く「恥ずかしい」が少ない：京都，大阪

「好き」「残す」「恥ずかしい」すべて多い：青森，岩手，宮城，秋田，山形，福島，島根，徳島，鹿児島

「好き」「残す」が多く「恥ずかしい」は平均並：北海道，長野，高知，福岡，長崎，熊本，宮崎，沖縄

《「土地のことば」に愛着がないグループ》

「好き」「残す」「恥ずかしい」すべて少ない：千葉，奈良，埼玉，神奈川

「好き」「残す」が少なく「恥ずかしい」が多い：茨城，栃木，岐阜，和歌山，岡山

「好き」「残す」が少なく「恥ずかしい」は平均並：滋賀

《すべて平均的なグループ》

「好き」「残す」「恥ずかしい」すべて平均並：静岡，三重，兵庫，鳥取，山口，香川，愛媛

《その他のグループ》

「その他」：群馬，東京，新潟，富山，石川，福井，山梨，愛知，広島，佐賀，大分

【類型⑥（田中 2011）】

《共通語中心社会群》

「首都圏・北海道型」：「方言」と「共通語」との区別意識をもたない。それゆえ、「方言」と「共通語」の使い分け意識も低い。「共通語」で通す意識が強い。

「北関東・甲信越型」：全体として「首都圏・北海道型」と同じような傾向。ただし、「どちらでもない」といった中間的な回答傾向を示す。

《方言主流社会群》

「近畿型」：「異郷友人」に対しても「出身地方言」の使用率が高い。「出身地方言」が好きで「共通語」の使用率・好感度とも低い。「方言」と「共通語」の使い分け意識は低く、「方言」で通す意識が強い。

「沖縄型」：「家族」や「同郷友人」に対する「出身地方言」使用率は高い。「出身地方言」も「共

「通語」もともに非常に好き。「方言」と「共通語」の使い分け意識も極めて高い。

「中国・九州・四国型」：「出身地方言」が非常に好きで、「家族」「同郷友人」に対する「出身地方言」使用率が高い。「共通語」はあまり好きではないが、その使用率も「方言」と「共通語」の使い分け意識も高い。

「北陸・東海・東北型」：「出身地方言」を「同郷友人」に対して「使うことがある」という遠慮がちな「出身地方言」使用パターン。共通語の「好き」はさほど高くないが、「方言」と「共通語」の使い分け意識は高い。

1.2.3 先行研究間の異同

以上、「方言」と「共通語」にかんする意識から導き出された地域や話者の類型を六つの先行研究を上げてみてきた。全体としてみると、調査年や分類対象地域の範囲による違いや、質問による違い、など細かな異同はあるものの、タイプ数の異なりを除くとそれほど大きな差異はないことがわかる。

タイプ数は最小3、最大8で、研究者個人の内省・観察に基づく類型にタイプ数が少なく、大規模データに基づき帰納的に類型を求めたものにタイプ数が多い。内省・観察に基づくものは、明確な特徴を示すものが類型として抽出されやすいためと推測される。大規模データに基づくものにタイプ数が多い理由は、どのタイプにも属さない中間的なふるまいをする話者・地域や、典型に準ずるタイプが帰納的に分離されることによると考えられる。帰納的手法に基づく類型化によって内省・観察では埋没していた輪郭のはっきりしないタイプが存在することが明示的になってきたことになる。輪郭のはっきりしないタイプも現実には存在しているわけで、大規模データによる帰納的手法から導かれる類型は、内省・観察のような直感的方法では捉えにくいものを抽出することができるという利点があるといえるだろう。

1.3 潜在クラスアプローチについて

1.3.1 潜在クラスモデルの目的

本稿では方言使用にかんする社会調査データに基づいて、個人の類型化を目指した統計手法として、潜在クラス分析（たとえば McCutcheon 1987, Hagenaars and McCutcheon 2002 を参照）を用いる。潜在クラス分析とは、いわゆるカテゴリカルな多変量データに対し「潜在クラス」（他にクラスタなどいくつかの呼び名がある）と呼ばれるカテゴリカルな潜在変数を導入することによって、変数間の連関を説明することを目指した分析手法である。対象者×分析対象変数のような調査データの型に即した用語で説明すると、潜在クラス所与の下での対象者の複数項目に対する応答は（条件付き）独立であることを仮定する。言い換えると潜在クラスが変数間に見られる連関の原因となっていることを仮定している。量的変数に対する分析手法である因子分析は変数間の相関関係の原因となるような潜在変数を導入するモデルであり、潜在クラス分析はそのカテゴリカル変数版であるといったたとえも成立する。

別の見方をすると、反応傾向の異なる複数の異質なグループの混合によって全体が得られると

いう Mixture Model (McLachlan and Peel 2000) の一種でもあり、潜在クラス分析は、各調査項目に対応する多項分布の Mixture を考える分析手法である。対象者のクラスタリングに用いることができるが、これに確率的な表現を与えるので、Probabilistic なクラスタリングの手法とみなされる。

たとえばマーケティング分野では、消費者は異質な嗜好・消費性向を有する複数のグループから構成されていると考える。そのような複数の異質な（内部では同質の）グループはセグメントと呼ばれ、データからセグメントを見つけることを「セグメンテーション」という。潜在クラス分析はセグメンテーションのための（もちろん唯一のということではないが）有力な方法である（岡太・守口 2010）。

本稿では、方言使用にかんする日本人の「セグメンテーション」を求めるこことを意図しており、手法の性格からこれは Probabilistic な方言話者の類型を措定することに他ならない。田中（2011: 92–114）ではこうした類型化を、地域ごとの集計（マクロ）レベルのデータに基づく考察によって達成したが、本稿の方法は個人（ミクロ）レベルの測定対象を直接（確率的に）類型化することを本旨とする分析モデルである点に留意する。日本全体を千数百人の調査データで代表させるためには方言はおそらく多様すぎる現象であるため、この類型化は大括りなものにならざるを得ず、各類型の中の小類型は捉えきれるわけではない。しかし、従来の検討との比較に用いることはでき、とくに両者（集計レベルでの考察と個人レベルでの考察）が同一の結論を導くのか、異なる結論を導くとしたらその原因は何か、という点が論点となりえるだろう。

1.3.2 潜在クラス分析の考え方

ここで潜在クラス分析のキー概念となる（潜在クラス所与の下での）条件付き独立の考え方のみ簡単に説明する。この部分が結果の解釈で利用される数値の説明となるからである。潜在クラスの定式化には大きく二つの方式がある。一つは、個々の質的な変数のカテゴリに対する反応確率を直接用いた表現で、これらの反応確率を複数の変数についてまとめた表（プロファイル）は結果の解釈に用いられる。もう一方の表現は、名義的な各反応変数の各カテゴリへの応答を多項ロジットモデルで表現する定式化である。前者は後者から再表現可能なものであるから、推定値の計算など統計的な取り扱いが容易な後者の定式化によってモデルを表現する場合も多い。他方、潜在クラスの解釈を行なうなど実質的な考察には、意味の理解が容易な前者の表現を用いることが普通である。なお潜在クラス分析には、反応変数が順序尺度水準の変数いわゆる順序付きカテゴリカル変数（ordered categorical variables）であることを想定するなどのバリエーションがあり得るが、本稿で利用したのは素朴にすべての反応変数が名義水準のカテゴリカル変数であることを想定したモデルである。

ここでは本稿で結果の表示にプロファイルを用いることに合わせて反応確率による定式化により、潜在クラス所与の下での反応の（条件付き）独立の意味を簡単に説明する。たとえば項目 $Y_{1(i)}$, $Y_{2(j)}$, $Y_{3(k)}$ の三つの質的変数（カテゴリ数はそれぞれ I, J, K 個）があり、 $Y_{1(i)}=1, 2, \dots, i, \dots, I$ のように値をとる。個々の変数の反応確率を π_i , π_j , π_k のように簡潔に表すことにする。 π_i は $Y_{1(i)}$ が

カテゴリ i の値をとる周辺確率の意である。また三つの変数の値が (i, j, k) となる同時反応確率を π_{ijk} のように表す。潜在クラス X は $(X=1, 2, \dots, C)$ のように C 個のカテゴリをもつ質的潜在変数として表現される。潜在クラス所与の下での各反応確率を $\pi_{i|X}, \pi_{j|X}, \pi_{k|X}, \pi_{ijk|X}$ のように表すと、先に述べた条件付き独立の仮定により、潜在クラスモデルでは同時反応確率が周辺反応確率の積で表現されることになる：

$$\pi_{ijk|X} = \pi_{i|X} \times \pi_{j|X} \times \pi_{k|X} \quad (i=1, 2, \dots, I, j=1, 2, \dots, J, k=1, 2, \dots, K, X=1, 2, \dots, C)$$

これは潜在クラス X のみが $Y_{1(i)}, Y_{2(j)}, Y_{3(k)}$ 間の連関の要因になっていることと潜在クラス内ではメンバーが同質であることを意味する。 $\pi_{i|X}, \pi_{j|X}, \pi_{k|X}$ を X の値(クラス)ごとに並べた値(プロファイル)を見て、各クラスの性格を考察することになる。

以上は、分析に利用する変数が、反応変数だけであるもっとも単純なモデルであった。本稿で実際に用いたモデルは、さらに潜在クラスを一つの名義的な応答変数とみなし、潜在クラス自身も他の説明変数(共変量)によって説明される被説明変数となっているモデル(共変量を伴う潜在クラスモデル(Dayton and Macready 1988, Vermunt and Magidson 2005))である。潜在クラス自体が被説明変数となる部分は、技術的には多項ロジットモデルで定式化される。

このように説明変数を導入することによって、本稿が目指す方言使用の個人類型を得るだけではなく、どのような個人の背景要因・社会的要因が規定力をもつかを考察することができる。

1.4 本稿の目的と構成

1.2 節までに述べた背景・先行研究を踏まえて、本稿の目的を次の点に置く。

第一に方言使用にかんする個人レベルでの類型(確率的な話者分類)を抽出すること、その類型と先行研究での類型の整合性を考察すること。

第二に上記の類型への帰属に対して個人の背景要因や社会的要因が与える影響を考察すること。以上の二点が主目的であるが、第二点の検討に際し、方言使用にあたって決定的と思われる(生育地にかんする)地域性と年齢要因について、やや詳しく考察する。とくに潜在クラス分析による個人のレベルの話者分類について、地域との対応関係との検証により地域類型としての位置づけを与えることを目指す。

この先、本稿は以下のように構成される。まず第2.1節で、分析に使用する調査データ、分析対象の変数とその再カテゴリ化の概要を説明する。続いて2.2節では分析モデルを説明する。

3節は全体が分析結果の説明である。3.1節では潜在クラス分析の結果の主要部分を、反応変数のプロファイルによって説明し、クラスの解釈(クラスのニックネームの付与)を行なう。続いて説明変数がクラス帰属に与える影響についても考察する。3.2節では潜在クラスと他の様々な(説明)変数との関連分析とくに帰属確率の記述的統計分析による各クラスの特徴づけをより詳細に行い、とくに地域と年齢変数がクラス帰属に与える影響を考察する。3.3節では地域と年齢が方言使用の類型に与える影響について、主要な五つの生育地域ごとに年齢の効果を組み合わせて検討する。

4節では、3節で得られたクラスによる地域類型に対する詳細な検討を行い、先行研究で得ら

れている類型との対応関係にかんする考察も行なう。5節では、本稿で残された課題について簡単に言及する。

2. 方法

2.1 2010年全国方言意識調査のデータ概要

本稿の分析で用いるデータは、層化三段無作為抽出法で抽出した16歳以上の全国の男女4,190人を対象に実施した「方言」と「共通語」についての意識を尋ねた全国方言意識調査に基づくものである。2010年12月に実施したもので、以下ではこの調査のことを「2010年全国方言意識調査」とする。

調査企画は筆者らが行ない、調査の実施は調査会社（中央調査社）に委託した。調査方法は、調査会社調査員による個別面接聴取法で、1,347人から回答が得られた（回収率32.1%）。なお、この全国方言意識調査は、筆者が共同研究者として参加している国立国語研究所基幹型共同研究「多角的アプローチによる現代日本語の動態の解明（プロジェクトリーダー：相澤正夫）」の一環として実施したものである。

47都道府県すべての居住者から回答を得ているが、本稿では、居住地とは別に「出身地」として質問した15歳までに一番長く生活した地域を「生育地」とみなし、生育地ごとの比較を地域差としてみていく。生育地は、下記に示す13の《生育地ブロック》レベルで検討する。なお、生育地を「海外」「わからない」とした6人については、「その他・不明」としてまとめた。

《生育地ブロック》

北海道（66人）、東北（128人）、北関東（85人）、首都圏（273人）、甲信越（72人）、北陸（44人）、東海（144人）、近畿（198人）、中国（89人）、四国（53人）、九州（169人）、沖縄（20人）、その他・不明（海外4人、わからない2人）※首都圏：東京都・埼玉県・千葉県・神奈川県

分析に用いる項目は、先に示した《生育地ブロック》を含む回答者属性6項目（性、年齢、職業、教育、居住地都市規模）と、「方言」と「共通語」にかんする質問7項目である。質問7項目の質問文を以下に示す。用いた選択肢は〔 〕内に示す。

[問3] あなたは、出身地（=15歳までに一番長く生活した場所）の「方言」のことが好きですか、嫌いですか〔好き、どちらかというと好き、どちらでもない、どちらかというと嫌い、嫌い、わからない〕

[問4] ここにあげる（A）～（C）の相手（（A）家族、（B）同じ出身地の友人、（C）異なる出身地の友人）に対して、出身地（=15歳までに一番長く生活した場所）の「方言」を使うことがありますか。それぞれ当てはまるものを一つずつ選んでください〔よく使う、使うことがある、使わない、わからない〕

[問7] あなたは、ふだんの生活において「共通語」を使っていると思いますか、思いませんか〔使っ

ていると思う、使っていると思わない、わからない]

[問8] あなたは、「方言」と「共通語」を場面によって使い分けていると思いますか、思いませんか〔使い分けていると思う、使い分けていないと思う、わからない〕

[問9] あなたは、「共通語」のことが好きですか、嫌いですか〔好き、どちらかというと好き、どちらでもない、どちらかというと嫌い、嫌い、わからない〕

なお、「方言」と「共通語」に対する好悪についての回答は、いずれも傾向性をはっきりさせるために「好き（好き+どちらかというと好き）」「どちらでもない」「嫌い（嫌い+どちらかというと嫌い）」と再カテゴリ化したものを、以下の分析において用いる。

生育地の「方言」使用にかんする質問は、「(A) 家族」「(B) 同郷友人（以下、同郷友人）」「(C) 異なる出身地の友人（以下、異郷友人）」といいういすれも親密な間柄の相手とのやりとりに限定したものである。このような私的場面においては、「方言」使用に傾くことが先行研究から明らかとなっている。つまり、2010年全国方言意識調査における「方言」使用にかんする設問は、私的場面に焦点を絞ったものとなっている。

2.2 分析の方法

1.3節で述べたように本稿で利用した分析モデルは共変量（説明変数）を伴う潜在クラス分析とよばれるものである。調査地域ごとに回収率が多少異なり、回収標本をそのまま分析すると、回収率が高かった地域の方言使用行動を母集団全体の中では過大評価する（回収率が低かった地域を過小評価する）ことにつながる。これは各類型のサイズの推定値にとくに影響を与えることが考えられる。この問題に対処するため、事後層化（Cochran 1977）の考え方に基づく補正ウェイトをつけた分析を行なった²。

利用した項目（内容は2.1節参照）を潜在クラスによって説明される反応変数（指標）と潜在クラスに対する説明変数として導入した属性項目に分けて再掲すると、次の通り。

潜在クラスの反応変数：「共通語と方言」にかんする7項目

問3：出身地の方言好悪、問4：出身地の方言を使うこと（A：家族、B：同じ出身地の友人、C：異なる出身地の友人、の3項目）、問7：ふだんの共通語使用、問8：方言と共通語の使い分け、問9：共通語好悪

共変量としてクラスに対する説明要因に導入した属性項目（6項目）

生育地ブロック、性、年齢、職業、教育、居住地都市規模

他に重要な属性要因として、生育地ではなく対象者自身が現在居住し（調査を受けた）居住地の

² ただし、層として設定された地域ブロックごとの回収率は30.2%から36.3%と大きな差はなく、ウェイトなしの分析との比較では、クラスの解釈に影響があるような本質的な差には至らなかった。以下ではウェイト付きの分析結果のみを示す。

地域性が考えられるが、生育地ブロックとの情報の重複が多く、同時に説明要因として取り上げるのは適切ではないので、モデルを推定後に事後的に考察するにとどめる。

本稿の分析では潜在クラスの推定には Latent Gold (Vermunt and Magidson 2005) を、それ以外の分析には SPSS Ver.17 を用いた。

3. 結果

3.1 潜在クラス分析の結果

潜在クラス分析を実行する上で、一般に分析者が最も判断に迷う点は、クラス数の決定である。このあたりの事情は因子分析における因子数の選択と同様に、決定的な手段があるわけではない。情報量規準 AIC や BIC などの指標を参考しながら、解釈可能性にも配慮しつつクラス数を五つに設定した解がもっとも適切であると判断した。この数までクラスを増やしていくと、情報量規準の意味で指標は確実に改善される一方、それよりもクラス数を大きくすると、クラスのサイズが 1,2% のようなマイナーなクラスが分離されることや、それらの妥当な解釈も見当たらないことから、1,300 程度という本サンプルから精度よく取り出せる情報は 5 クラスの解が限度であると判断したものである。

表 1 は（共変量を伴う）潜在クラスモデルの主要部分の結果を（反応変数に用いた 6 項目に対する各クラスの応答プロファイルとして表現したものである。

このプロファイルを参照すると、各クラスは次のように解釈する（クラスのニックネームをつける）ことができる。クラス名のあとに（ ）内にはクラスの相対的大きさ（クラスサイズ）を転記しており、その下の説明文は六つの項目に対するプロファイルを言語表現したものである。

クラス 1：積極的方言話者（0.373）

生育地の方言が好きと答える確率が高く（問 3）、家族、同郷の友人、異郷の友人のいずれに対しても生育地の方言をよく使っている（問 4）。ふだん共通語を使っているという認識が五つクラス中でもっとも低く（問 7）、共通語と方言の使い分け意識はそれほど高くなく（問 8）、共通語への好意度は五つのクラスでもっとも低い（問 9）。

クラス 2：共通語話者（0.286）

誰に対してもふだん生育地の方言を使うことが少なく（問 4）、方言との使い分けというよりも共通語を使っていると認識しており（問 8 と問 7）、共通語への好意度も高め（問 9）である。

クラス 3：消極的使い分け派（0.159）

クラス 1 やクラス 4 のような方言使用者に比べると、生育地の方言への好意度がやや低く（問 3）、生育地の方言は家族や同郷の友人に対してやや頻度が高く、異郷の友人に対しては低いもの、全体に「使うことがある」程度（問 4）。共通語と方言の使い分け意識が高いわけではなく（問 8）、ふだんの共通語使用の程度（問 7）はクラス 2 やクラス 4 に比べれば低い。共通語に対する好意

表1 7つの反応変数（調査項目）に関するプロファイル

反応変数とそのカテゴリ	クラス 1	クラス 2	クラス 3	クラス 4	クラス 5	全体
	積極的方言 話者	共通語 話者	消極的使い 分け派	積極的使い 分け派	判断逡巡派	
クラスのサイズ	0.373	0.286	0.159	0.149	0.033	1.000
問3：出身地の方言好悪						
好き	0.786	0.384	0.535	0.851	0.123	0.619
どちらでもない	0.184	0.461	0.389	0.063	0.621	0.292
嫌い	0.027	0.067	0.062	0.081	0.072	0.053
わからない	0.002	0.089	0.014	0.005	0.184	0.035
問4a：出身地の方言を使うこと—家族						
よく使う	0.987	0.006	0.022	0.351	0.044	0.427
使うことがある	0.003	0.030	0.935	0.390	0.072	0.219
使わない	0.006	0.962	0.044	0.254	0.189	0.328
わからない	0.004	0.002	0.000	0.005	0.695	0.026
問4b：出身地の方言を使うこと—同じ出身地の友人						
よく使う	0.992	0.000	0.032	0.500	0.024	0.451
使うことがある	0.000	0.039	0.968	0.395	0.001	0.224
使わない	0.000	0.958	0.001	0.089	0.002	0.288
わからない	0.008	0.003	0.000	0.016	0.973	0.038
問4c：出身地の方言を使うこと—異なる出身地の友人						
よく使う	0.706	0.003	0.006	0.023	0.002	0.269
使うことがある	0.207	0.003	0.674	0.295	0.001	0.229
使わない	0.079	0.978	0.266	0.655	0.021	0.450
わからない	0.007	0.017	0.054	0.028	0.977	0.052
問7：ふだんの共通語使用						
使っていると思う	0.555	0.858	0.689	0.888	0.676	0.717
使っていると思わない	0.412	0.093	0.236	0.105	0.118	0.238
わからない	0.033	0.049	0.076	0.006	0.206	0.046
問8：方言と共通語の使い分け						
使い分けていると思う	0.410	0.142	0.348	0.836	0.118	0.378
使い分けていないと思う	0.539	0.767	0.540	0.164	0.538	0.548
わからない	0.051	0.091	0.112	0.000	0.344	0.074
問9：共通語好悪						
好き	0.467	0.638	0.420	0.681	0.404	0.538
どちらでもない	0.474	0.330	0.528	0.294	0.411	0.412
嫌い	0.053	0.017	0.030	0.000	0.000	0.029
わからない	0.007	0.016	0.021	0.024	0.186	0.020

注) 表は、項目ごとに列（縦方向）の和が1（100%）となる。

度（問9）は「どちらでもない」が多数派である。方言の使用頻度はある程度高いが、強く使い分けを意識しているわけではなく、共通語に対するややアンビバレン特な態度がうかがわれる。

クラス4：積極的使い分け派（0.149）

生育地の方言が好きと答える確率がもっとも高いクラス（問3）で、家族や同郷の友人には生育地の方言を使うことがあるが、異郷の友人には方言は使わない（問4）。ふだん共通語を使って

いるという意識も高く（問7），それは方言と共通語の使い分け意識の高さ（問8）にも現われている。共通語に対する好意度（問9）ももっとも高い部類のクラスである。

クラス5：判断逡巡派（0.033）

このクラスは五つのクラスの中では相対的な重みが小さい少数派であるが，このクラスの特徴は多くの間にわからないと答える傾向がある点といえる。ふだんの共通語使用（問7）や使い分け意識（問8）について尋ねられても，他のクラスに比べるとわからないという回答の確率が高くなっている，方言使用について尋ねられたときに判断に逡巡することを特徴とするクラスといえる。3.2, 3.3節で属性要因との関係を考察するが，この背景には二つの要因を指摘できる可能性がある。

以上のように応答変数に対するプロファイルによって五つのクラスの一応の意味づけが可能になった。

3.2 潜在クラスと帰属確率との記述的分析による各クラスの特徴づけ

以上のようなクラスの特徴は，説明変数との関係を調べることによってより明確化することができる。多項ロジット表現のパラメータである回帰係数のままでは数値の意味が捉えにくいので，その回帰係数を踏まえた上で，説明変数の各カテゴリにかんして，モデルの推定値に基づき五つの潜在クラスのそれぞれに帰属する事後確率の平均を表現した「クラス帰属確率」を求め，この指標を用いて関係を検討する。これは，ロジットモデルのパラメータそのものではないが，クラスの性格付けを考える上で有用な情報である。

ロジットモデルにおいて各説明変数の全体的效果にかんする統計的検定によれば，取り上げた説明変数の中で，クラス帰属への影響力が有意なのは，（効果が大きい順に）生育地，職業，教育，年代であり，居住地都市規模や性別は有意とはいえないかった。

表2には，取り上げたすべての共変量の各カテゴリについて，「クラス帰属確率」をまとめたものを示す。当該のカテゴリごとに五つのクラスへの帰属確率の和が1（つまり表の行方向の和が1）になるような数値である。

すでに述べたように，居住地都市規模や性別は説明変数として有意な効果を示していないので，表中のカテゴリ間の確率の差を積極的に解釈すべきではない。ここでは効果が有意であった説明変数のみに基づき，五つのクラスの属性面での特徴を述べる。クラスサイズとはこの帰属確率のいわば全体平均の値であるので，クラスサイズとの比較により，当該属性をもつ話者（回答者）の特徴をみることができる。クラスの特徴は次のように要約される。年齢の効果については全般にやや非線形な関係が認められるようであり，3.3節で詳しく検討する。

表2 説明変数のカテゴリ別帰属確率平均

	クラス 1 積極的方言 話者	クラス 2 共通語話者	クラス 3 消極的使い 分け派	クラス 4 積極的使い 分け派	クラス 5 判断逡巡派
クラスのサイズ	0.373	0.286	0.159	0.149	0.033
説明変数とそのカテゴリ					
性別					
男性	0.390	0.269	0.148	0.157	0.037
女性	0.359	0.301	0.169	0.142	0.029
年齢					
16–19 歳	0.344	0.336	0.124	0.145	0.051
20–29 歳	0.444	0.267	0.171	0.082	0.037
30–39 歳	0.400	0.251	0.134	0.200	0.015
40–49 歳	0.440	0.272	0.118	0.157	0.013
50–59 歳	0.261	0.336	0.182	0.184	0.036
60–69 歳	0.375	0.269	0.202	0.118	0.036
70 歳以上	0.332	0.304	0.164	0.149	0.052
職業					
農林漁業	0.677	0.224	0.100	0.000	0.000
商工・サービス業	0.401	0.293	0.163	0.115	0.028
事務職	0.373	0.253	0.146	0.189	0.039
労務職	0.405	0.249	0.146	0.187	0.013
自由業・管理職	0.318	0.219	0.191	0.272	0.000
無職の主婦	0.340	0.319	0.168	0.139	0.034
学生	0.352	0.362	0.124	0.130	0.033
その他無職	0.364	0.275	0.189	0.117	0.055
教育					
新制中学・旧制小学／高小	0.424	0.279	0.158	0.109	0.030
新制高校・旧制中学	0.375	0.284	0.165	0.139	0.036
新制短大／大学・旧制高専／大学	0.355	0.291	0.149	0.179	0.026
不明	0.000	0.329	0.329	0.000	0.343
生育地ブロック					
北海道	0.204	0.528	0.093	0.051	0.123
東北	0.301	0.245	0.152	0.256	0.048
北関東	0.202	0.256	0.359	0.183	0.000
首都圏	0.190	0.665	0.096	0.000	0.049
甲信越	0.214	0.240	0.337	0.196	0.014
北陸	0.345	0.095	0.342	0.218	0.000
東海	0.454	0.162	0.269	0.116	0.000
近畿	0.679	0.118	0.155	0.034	0.015
中国	0.503	0.090	0.103	0.248	0.057
四国	0.491	0.275	0.072	0.161	0.000
九州	0.433	0.111	0.073	0.354	0.029
沖縄	0.173	0.203	0.000	0.625	0.000
その他・不明	0.176	0.477	0.000	0.000	0.347
居住地都市規模					
20 大都市	0.343	0.343	0.146	0.109	0.059
市部	0.373	0.267	0.173	0.164	0.023
町村	0.447	0.254	0.110	0.168	0.021

注) 表は、行(横方向)の和が1(100%)となる。

クラス 1：積極的方言話者

生育地については近畿に属する対象者がこのクラスへの帰属確率が高い。ついで中国・四国の生育者が帰属しやすい。職業では農林漁業、教育は低めの層がこのクラスに帰属しやすい。近畿生育の関西弁話者に典型的にみられる方言使用類型とみられる。

クラス 2：共通語話者

生育地が首都圏・北海道と（調査地点の）都市規模が大きい対象者が帰属しやすい。北陸以西のいわゆる西日本では全般に帰属確率は低くなる。職業面では学生層と無職・主婦層の帰属確率が高めである。

クラス 3：消極的使い分け派

このクラスに対する帰属確率については生育地の特徴のみが顕著で、北関東・甲信越・北陸のような大都市圏周辺の地域の生育者が帰属しやすいクラスである。東海もこれらについて高めになっている。

クラス 4：積極的使い分け派

このクラスも地域的な特徴が顕著で、沖縄生育者が突出して高い帰属確率を示す。他に高いのは九州と東北ならびに中国の生育者であり、これらの地域の生育者は方言と共に語の使い分けを強く意識していることになる。逆に首都圏生育者はこのクラスに帰属確率がほとんどなく、近畿圏生育者もこのクラスには属しにくい。職業面では自由業・管理職にやや高めの値が観察される。

クラス 5：判断逡巡派

これは相対的にサイズの小さいクラスであるため、特徴がややつかみにくいが、地域的に北海道生育者にとくに高い点が特徴である。

以上が年齢を除く属性要因の効果を検討したものであるが、年齢については表2の10歳刻み幅での集計数値を参照すると効果が（年齢の上昇とともに）単調に増加したり減少したりするわけではなく、非線形な関係を予想させるものであるので、節を改めてやや詳細に検討する(3.3節)。

また参考として、分析モデルには導入しなかった対象者の居住地の効果を検討するために、事後的に居住地ブロックごとの帰属確率平均を求めたものについても、表3に示しておく。生育地のブロック区分と異なっており単純な対応関係にはないが、生育地の地域性にかんする特徴は、居住地についてもほぼ同様の傾向が確認される。たとえば近畿・阪神・中国・四国の居住者は「クラス1:積極的方言話者」に帰属する確率が高く、関東・京浜・北海道は「クラス2:共通語話者」に帰属する確率が高い、などである。

表3 居住地ブロック別の帰属確率平均

	クラス 1 積極的方言 話者	クラス 2 共通語話者	クラス 3 消極的使い 分け派	クラス 4 積極的使い 分け派	クラス 5 判断逡巡派
クラスのサイズ	0.373	0.286	0.159	0.149	0.033
居住地ブロック					
北海道	0.228	0.525	0.061	0.046	0.141
東北	0.358	0.197	0.164	0.229	0.051
関東	0.193	0.495	0.175	0.121	0.016
京浜	0.180	0.561	0.098	0.089	0.072
甲信越	0.221	0.232	0.360	0.172	0.015
北陸	0.365	0.091	0.364	0.181	0.000
東海	0.429	0.186	0.259	0.120	0.006
近畿	0.622	0.126	0.137	0.107	0.009
阪神	0.640	0.118	0.146	0.066	0.031
中国	0.544	0.112	0.120	0.163	0.062
四国	0.527	0.251	0.104	0.118	0.000
九州	0.478	0.084	0.062	0.349	0.027

3.3 年齢効果の詳細な吟味

ここでは各クラスへの帰属確率への年齢間の相違を視覚的に検討しておくこととする。

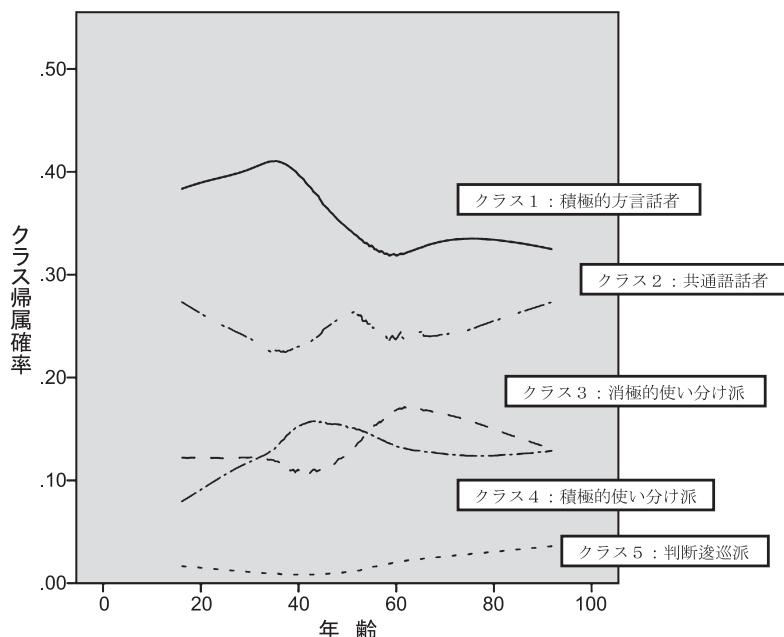


図1 年齢別クラス帰属確率平均値（平滑化後の曲線）

図1は個人の五つのクラスに対する帰属確率（縦軸）を年齢（横軸）に対する散布図に、局所線形化回帰と呼ばれるノンパラメトリックな方法（Wand and Jones 1997）で各年齢の平均値をスマージングした曲線を書き入れたものである（ただし個人のプロットを入れると見にくくなるので、プロットは表示していない）。曲線は対象者年齢の最大値（92歳）まで表示されているが、このような最高齢層の部分はサンプルサイズが小さいので、結果は不安定と考えるべきである。

図1からは、以下のようなことが読み取れる。

「クラス1：積極的方言話者」への帰属確率は、40歳前くらいの年齢で最大となり、60歳くらいまでの20歳ほどの年齢層ではほぼ単調に減少、60歳より上の年齢層で多少の持ち直しがみられる。

他方、「クラス2：共通語話者」への帰属確率の年齢差の様相はやや複雑で、50歳前後に一つのピークがあり、35歳と60歳手前ほどの箇所に小さな底がある逆M字型の様相を示している。50歳前後を境に、共通語に対する態度の差がみられる可能性を指摘できるかもしれない。

「クラス3：消極的使い分け派」は40歳が底で60歳が最大になる。クラス1の積極的方言話者の年齢差の裏返しの形になっている。

「クラス4：積極的使い分け派」は、40代前半をピークとする山型になっている。

「クラス5：判断逡巡派」は、40代前半くらいを底にそれより上の年齢層でゆるやかに上昇していくが、一般に社会調査では高齢層にわからない回答が増える傾向があり、クラス5の結果は、北海道という地域性の他、年齢層が高くなるのにしたがって、方言使用の実態とは無関係に調査の設問に対する回答として全般に「わからない」という「自信のない」回答をしがちであることを反映しているようである。

ここで述べたクラスへの帰属確率の年齢間の差にかんする解釈は、地域性を無視して全国的に見た場合の傾向であって、年齢の効果だけに単純な解釈を与えることは危険である。先に見たようにクラスへの帰属確率については生育地域の効果が突出して大きく、その地域性も考慮した上での分析が望ましいが、本稿の調査データではすべての地域ごとに年齢の効果を細かく検討するのに十分なサンプルサイズはない。ここでは層別に解析しても比較的大きなサンプルサイズが確保できる近畿（n=198）、首都圏（n=273）、東海（n=144）、九州（n=169）、東北（n=128）の5地域についてのみ、年齢にかんする平滑化された帰属確率平均値を示し（図2a～e）、地域ごとに年齢の特徴を考察する。この5地域についてもサンプルサイズが小さいための結果の不安定さが伴い、記述内容は本稿での暫定的な考察にとどまる。また地域によっては確率が0に近い対象者が多いクラスの帰属確率は図に示せなくなっている（図2bのクラス4と図2cのクラス5）。

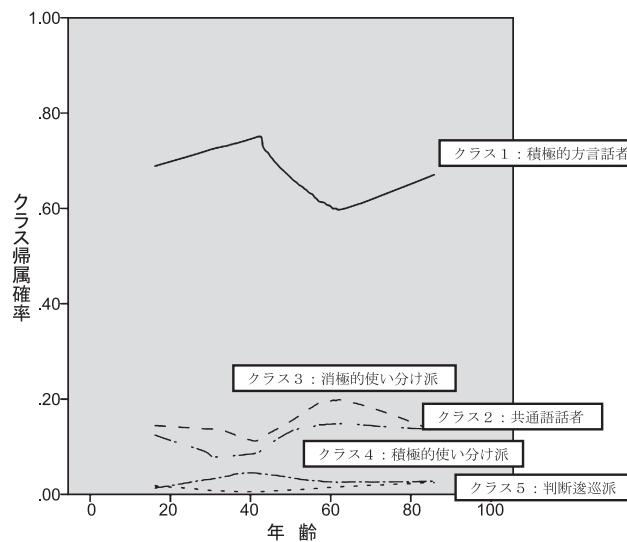


図 2a. 近畿圏生育者の年齢別クラス帰属確率

たとえば近畿圏生育者（図 2a）では、「クラス 1：積極的方言話者」に対する帰属確率がもっとも高いが、年齢の効果についてはピークである 40 代前半くらいまでゆるやかに高くなつた後、60 代前半の底に向けて減少し、その後再度上昇するパターンがみられる。

首都圏生育者（図 2b）は「クラス 2：共通語話者」が最大クラスであるが、30 歳前くらいが底、50 代後半がピークでその後また減少するパターンを示している。

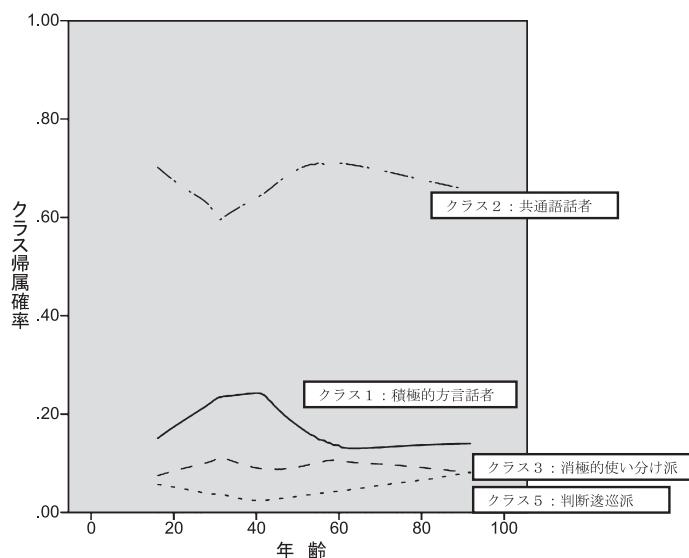


図 2b. 首都圏生育者の年齢別クラス帰属確率

東海生育者（図2c）は「クラス1：積極的方言話者」が最大クラスで、近畿生育者とほぼ類似のパターンを示す。続いて多いのは「クラス3：消極的使い分け派」で、40代から60歳くらいにかけての上昇も観察される。

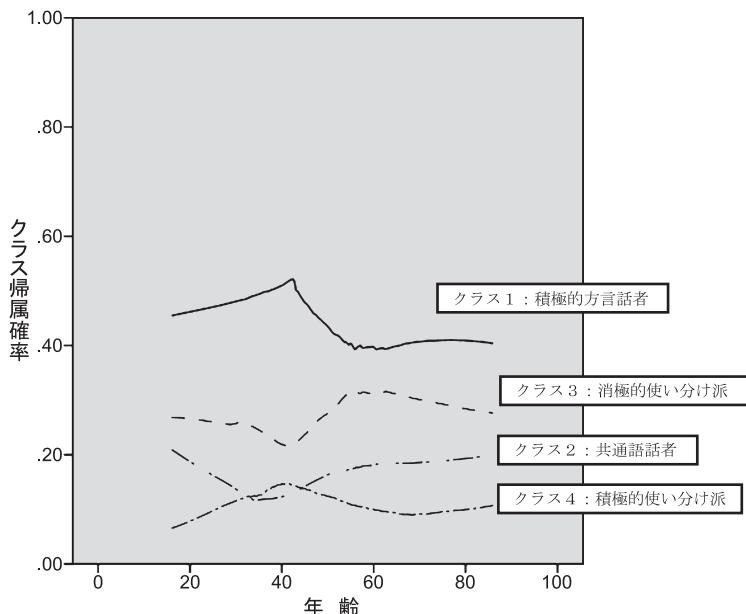


図2c. 東海生育者の年齢別クラス帰属確率

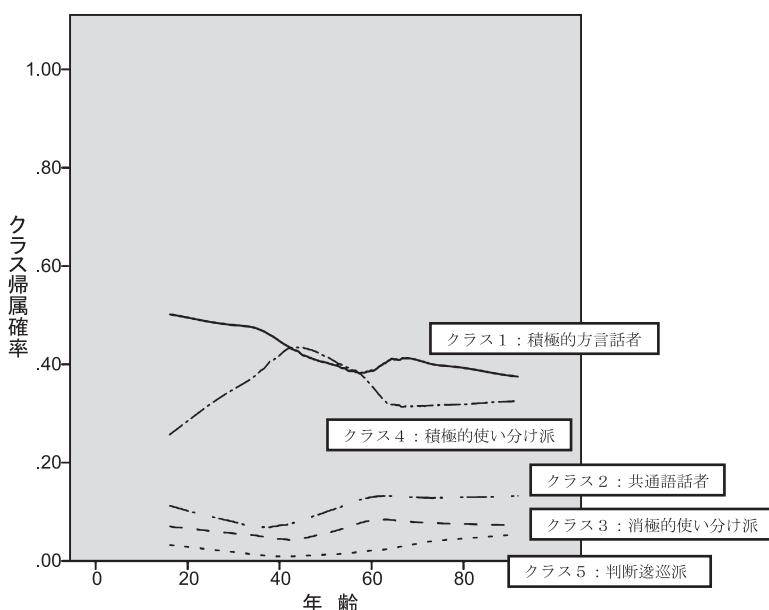


図2d. 九州生育者の年齢別クラス帰属確率

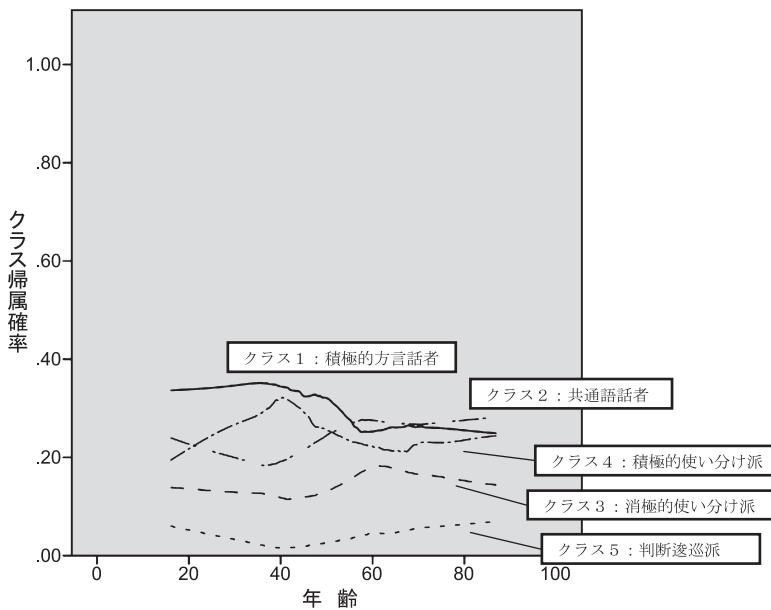


図 2e. 東北生育者の年齢別クラス帰属確率

九州生育者（図 2d）は「クラス 1：積極的方言話者」と「クラス 4：積極的使い分け派」が相対的に大きなクラスであるが、前者は年齢とともに減少基調、後者は若年層から 40 歳くらいまで上昇基調でその後下降した後、60 代以降で下げ止まりかむしろやや高まる傾向も示している。

東北生育者（図 2e）は、他の地方にくらべてクラス間のサイズの違いが小さいことがまず特徴になっている。最大は「クラス 1：積極的方言話者」であるが、40 歳あたりを境に減少、ついで「クラス 4：積極的使い分け派」と「クラス 2：共通語話者」が拮抗するが、クラス 2 は二つの方言使用クラスの減少を補うように 40 歳あたりから上昇に転じて 60 歳あたりで高止まりとなる。

これらの年齢傾向については、4.3 節で改めて議論する。

4. 考察

ここでは、2010 年全国方言意識調査データを用いた話者の分類から五つの潜在クラスが抽出された意義について検討を試みる。

まず、五つの潜在クラスを帰属確率の高い回答者生育地との関係から地域類型としてみなし、どのような背景から抽出された地域類型であるのかについて考察する。次に、今回の試行から導き出された地域類型が、先行研究において示された地域類型とはどのような対応関係にあるのかを検討する。最後に、3.3 節でみてきたデータ全体における各クラスへの年齢による帰属確率の変化と、層別分析に耐えうるサンプル数が確保された近畿・首都圏・東海・九州・東北の五つの生育地群ごとの年齢によるクラス帰属確率の変化の示す意味について考察を行なう。

4.1 五つの潜在クラスから導き出される地域類型

五つの潜在クラスと各クラスにおける回答者生育地帰属確率が相対的に高い群との対応関係は、3.2節でみてきた通り。ここでは、それぞれの生育地群が各クラスに高い帰属確率を示す背景を検討するとともに、各クラスの典型を抽出し、「方言」と「共通語」に対する言語意識に基づく地域類型を導き出していきたい。

まず、「クラス1：積極的方言話者」への帰属確率の高い生育地群は、近畿を中心とした西日本域の方言主流社会であることがわかる。近畿生育者の帰属確率が0.679ともっとも高く、クラス1の典型である。中国（帰属確率0.503）・四国（帰属確率0.491）が近畿とひとまとまりのものとして取り出された点が特徴的である。

次に、「クラス2：共通語話者」への帰属確率の高い生育地群は、こんにちの共通語基盤方言である首都圏（帰属確率0.665）と、北海道（帰属確率0.528）である。北海道方言は近代以降の移住者により形成されてきたという背景から、語彙の標準語形との一致率が首都圏並に高く（河西1981）、共通語との心理的距離も地理的距離に比して近い（言語編集部1995）。このようなことから共通語基盤方言である首都圏方言話者と言語意識が近く、地域類型として結びついたものと思われる。一方、「クラス5：判断逡巡派」はどの生育地群の帰属確率も高くないが、その中において北海道生育者の帰属確率が特徴的に高い（帰属確率0.123）という特徴が認められるのは、首都圏生育者ほど迷いなく自らの生育地方言を「共通語」と認識する水準ではないためであろう。ただし、クラス2とクラス5における北海道生育者の帰属確率を比較すると、クラス2への帰属確率が高く、北海道生育者の基本はクラス2、ついでクラス5となっている（表2）。このことは、北海道生育者は「共通語話者」という認識が主流であるが、ついで「判断逡巡派」が現われる、ということを示している。

続いて、「クラス3：消極的使い分け派」への帰属確率の高い生育地群は、北関東（帰属確率0.359）・甲信越（帰属確率0.337）・北陸（帰属確率0.342）・東海（帰属確率0.269）である。ただし、表2から生育地群ごとのもっとも帰属確率の高いクラスをみていくと、このうち北陸・東海生育者は「クラス1：積極的方言話者」への帰属確率がもっとも高い。クラス3に帰属する典型としては、他のクラスへの帰属確率が相対的に低い北関東と甲信越ということになる。いずれも首都圏の周辺域の生育地群という共通点がある。言語的には共通語基盤方言の周辺方言であり、政治・文化・経済的にも準首都圏的位置づけである地域が典型的に帰属するクラスとみることができる。

「クラス4：積極的使い分け派」への帰属確率の高い生育地群は、沖縄（帰属確率0.625）・九州（帰属確率0.354）・東北（帰属確率0.256）・中国（帰属確率0.248）である。ただし、表2からは、中国・九州・東北は「クラス1：積極的方言話者」への帰属確率がもっとも高いこともわかる。沖縄は、クラス4内の他生育地群との比較において帰属確率が群を抜いて高いばかりでなく、他のクラスへの帰属確率が低い。よってクラス4の典型は、沖縄とみることができる。沖縄方言は方言区画の観点からも他の本土方言との言語的距離のもっとも遠いものである（東條1954）。一方、共通語への心理的距離は、地理的距離・言語的距離に比して、非常に近い（言語編集部1995）。このような背景から、共通語への言語的距離の遠さと心理的距離の近さによって「クラス4：積極的

使い分け派」の典型として沖縄が抽出されてきたと考えられる。

「クラス3：消極的使い分け派」と「クラス4：積極的使い分け派」への帰属確率の高い生育地群については、それぞれの生育地方言がこんにちの日本語社会においてイメージ喚起力の強い、いわばはっきりした方言イメージのある地域か、そうでない地域か、という観点から説明可能だろう。友定（1999）、田中（2011: 70–80）などから、クラス4への帰属確率の高い沖縄・九州・東北・中国は、イメージ喚起力の強い地域・都市を含むが、クラス3への帰属確率の高い北関東・甲信越・北陸・東海はイメージ喚起力の強い地域が含まれない、という共通点があることがわかる³。

以上から、2010年全国方言意識調査データを用いた話者の分類に基づく地域類型を次のように示すことができるだろう（□内の生育地群はクラス内・生育地群内の帰属確率の比較からみた典型）。

「クラス1：積極的方言話者」：近畿・中国・四国

「クラス2：共通語話者」：首都圏・北海道

「クラス3：消極的使い分け派」：北関東・甲信越・北陸・東海

「クラス4：積極的使い分け派」：沖縄・九州・東北・中国

「クラス5：判断送巡派」：北海道

4.2 先行研究における地域類型との比較

ここでは、4.1節において導き出した「方言」と「共通語」にかんする言語意識による地域類型を、1.2節で概観した先行研究における地域類型と比較する。4.1節で導いた地域類型と、先行研究の類型とを対応づけたものが表4（陣内1996、1999の類型と地域の対応については、1.2節参照のこと）。ただし、表4では、移住地における言語使用の心理的傾向性の類型を示している野林（1971）と、本稿筆者によるデータの単純な分類に留まっているNHK放送文化研究所編（1997）については、検討の対象としにくいため、比較の対象から外してある。

表4 五つの潜在クラスと先行研究の類型との対応

クラス:ニックネーム[帰属確率の高い生育地群:太字は典型]	寿岳(1978)	陣内(1996)	陣内(1999)	田中(2011)
クラス1:積極的方言話者[近畿・中国・四国]	関西型	無意識顕在化型	方言開示型	近畿型
クラス2:共通語話者[首都圏・北海道]	沖縄型	—	(方言開示型)	首都圏・北海道型
クラス3:消極的使い分け派[北関東・甲信越・北陸・東海]	東北型			北陸・東海・東北型
クラス4:積極的使い分け派[沖縄・九州・東北・中国]		(意識的潜在化型)(中間型)	使い分け型	沖縄型 中国・九州・四国型
クラス5:判断送巡派[北海道]	—		(中間型)	北関東・甲信越型
潜在クラスとの対応なし	—	(意識的潜在化型)(中間型)	方言抑制型 (中間型)	—

[注] —は対応する類型がないことを表し、()はストレートな対応関係の有無の判断に迷うものを表す。

³ 友定（1999）では、全国14都市の方言をイメージ喚起力の観点から「イメージ濃厚方言」と「イメージ希薄方言」に二分している。「弘前」「京都」「広島」「福岡」「那覇」が「イメージ濃厚方言」に分類されている。また、田中（2011: 74）では、2007年と2010年に実施した2種類の方言イメージ調査に共通するイメージ喚起力の強い方言として、「青森」「東京」「京都」「大阪」「広島」「福岡」「熊本」「鹿児島」「沖縄」を指摘している。いずれにおいても、沖縄・九州・東北・中国に位置する地域が含まれていることがわかる。

表4から、寿岳（1978）以降の地域類型において、一貫して相手・場所を問わずに生育地方言の使用に傾く「方言開示型」「積極的方言話者」の類型に配置される地域として近畿を指摘できる。

東北も、寿岳（1978）以降、一貫して地元「方言」・非地元「共通語」という使い分けを行なう類型に位置づけられていることは共通する。しかし、使い分けの態度による区分が抽出された田中（2011）では、「出身地方言を「同郷友人」に対して「使うことがある」という遠慮がちな」タイプに分類されたが、同じデータに対する話者レベルのミクロな分類から類型を抽出した4.1節では、「クラス4：積極的使い分け派」に分類されるという差異が認められた。

一方、その他の地域は、類型化の対象に含まれることも含まれないこともあるのと同時に、類型としての位置づけにも異同があることがわかる。

類型間においてもっとも大きな異同が認められる地域は沖縄である。寿岳（1978）では「沖縄型〔地元でもよそでも共通語〕」という類型であったが、陣内（1999）では「方言抑制型〔どの場面でも方言使用が相対的に低い都市〕」に位置づけられている。4.1節と同じデータを用いた田中（2011）では地元「方言」・非地元「共通語」という使い分けがはっきりしており、「方言」も「共通語」も非常に好きという傾向を示す「沖縄型」として独立した類型を形成し、4.1節の類型においては「積極的使い分け派」の類型内に位置づけられている。これら研究間の類型の異同からは、沖縄の方言抑制から方言と共通語の積極的使い分け派へという大きな流れが確認できる。とくに田中（2011）と4.1節で導き出された本稿類型において明確化した沖縄の位置づけの変化は、2000年代以降に顕著となった日本語社会における沖縄文化や方言に対する好感度と沖縄県民の肯定感の上昇が関連すると思われる（多田2008⁴）。

寿岳（1978）で示されたような方言的特徴の明確な地域や方言に限定した類型化から、中間的・周辺的方言も含めた類型の抽出という方向も大きな変化として指摘することができる。同時に首都圏ならびに東京を類型の中に位置づけようという考え方も明確化してきていることがわかる。陣内（1996, 1999）では「中間型」が設けられ、田中（2011）ならびに4.1節では、それら中間型が「消極的使い分け派」「判断逡巡派」として区分できそうであることを示している。陣内（1999）では東京を地域分類の対象として取り上げ、「方言開示型〔どの場面でも相対的に方言使用率が高い都市〕」の一つに位置づけている。田中（2011）ならびに4.1節では一貫して「共通語」の使用率が高く、「方言」と「共通語」の使い分け意識が低い「共通語話者」類型に位置づけた。

以上、「方言」と「共通語」に対する使用意識からみた地域類型間の異同は、日本語社会における「方言」と「共通語」の関係性とそれに対応する価値ならびに各地域方言に対する意識の変化によると考えられる。それと同時に、地域類型間の異同は、年齢と各クラスへの帰属確率ならびに地域別の年齢による各クラス帰属確率の変動とより強く関わる問題もあるので、これらについては4.3節で詳しく検討していく。

⁴ 沖縄を舞台とした映画（『ナビィの恋』1999年公開、中江裕司監督）や、NHK連続テレビ小説（『ちゅらさん』2001年度前期放送）などを経て、沖縄の全国的な認知度が高まり、「県民に新たな承認と誇りの意識を与えた（多田2008）」としている。

4.3 年齢による各クラス帰属確率の変動

ここでは、3.3 節でみてきた各クラスへの年齢による帰属確率の変動の背景についてみていく。

まず、全体のデータを反映した図 1 から検討する。3.3 節で指摘した通り、60 歳前後と 40 歳前後が主要なクラスの主な帰属確率の変化点となっている。のちに検討する図 2a ~ e の生育地群別の図においてもほぼ 60 歳前後と 40 歳前後に帰属確率変化点が観察されることは共通している。

40 歳前後から 60 歳前後は社会的活躍層であるということから、ライフステージの観点において他の世代とは異なるふるまいをみせている可能性もあるが、地域によって帰属確率の変動パターンは少しずつ異なりつつも、変化点は共通するということからは、60 歳前後ならびに 40 歳前後の世代において、日本語社会全体の「方言」と「共通語」にかんする言語意識に影響を与える社会的变化があったことが推測される。

図 1 からはさらに、「クラス 1: 構築的方言話者」の帰属確率がもっとも高く、ついで「クラス 2: 共通語話者」の帰属確率が高いことは年齢を問わないが、60 歳から 40 歳前後の間にクラス 1 への帰属確率が大きく上昇していることがわかる。クラス 2 は 50 歳前後をピークとして 40 歳前後に向けてクラス 1 の帰属確率の上昇と反比例するように減少する。「クラス 3: 消極的使い分け派」と「クラス 4: 構築的使い分け派」は、50 歳より上の世代と 35 歳より下の世代においてはクラス 3 が優勢だが、50 歳前後から 35 歳前後においてはクラス 4 が優勢である。これらから、おおむね 60 歳前後から 40 歳前後までの世代においては、「クラス 1: 構築的方言話者」と「クラス 4: 構築的使い分け派」という「方言」を構築的に使用する意識が高いことがわかる。

60 歳前後と 40 歳前後の言語形成期は、それぞれ 1950 年代から 1960 年代と 1970 年代から 1980 年代である。前者は高度経済成長期、後者はバブル経済に向かって日本の経済ならびに社会構造が大きく変化した時期である。同時に、「方言」と「共通語」に対する認識が大きく変化しつつあった時期とも重なる。すなわち、「方言」が「恥ずかしい、かっこ悪い、隠す」ものから「誇らしい、かっこいい、見せる」ものへと日本語社会における位置づけが大きく変化しつつあった時期に、思春期・言語形成期を過ごした世代ということになる(田中 2011: 40–66)。こういった日本語社会全体における「方言」の価値の上昇といった変化が、クラス帰属確率の変化に影響を与えたと考えられる。もちろん、60 歳前後から 40 歳前後は社会的活躍層であるため、「共通語」を意識的に使用する場面も多く、「方言」との使い分けや、「方言」を方略的に用いる効果を意識しやすいため、前後の世代とは異なるふるまいをみせている、ということも否定できない。しかし、「方言」と「共通語」の意識的使い分けや、方略としての「方言」使用という意識が成立するためには、先に述べたような「方言」に対するネガティブなイメージがポジティブなものに移行しているということが前提となるだろう。

上述した図 1 から確認される傾向は、近畿（図 2a）・東海（図 2c）においてもほぼ同様に観察される。これらと並行的に、「クラス 2: 共通語話者」への帰属確率が全年齢を通じてもっとも高い首都圏（図 2b）においても、60 歳前後から 40 歳前後にかけてクラス 1 への帰属確率が急上昇している。共通語基盤方言域である首都圏生育者においても、「クラス 1: 構築的方言話者」

への帰属確率が上昇していることになる。これらのことから、60歳前後から40歳前後の世代にかけて「積極的方言話者」を志向するような「方言」に対する価値観の変化が日本語社会の中で進行してきたことが推測される。

一方、全体・近畿・東海・首都圏に共通して、40歳前後より若い世代においては、いずれも「クラス1: 積極的方言話者」への帰属確率が減少し、「クラス2: 共通語話者」への帰属確率が高まる。60歳前後から40歳前後にかけて進行していた「積極的方言話者」を志向する傾向が、反転していることになる。

この反転が示す意味については、少なくとも次の二つの解釈が成りたちそうである。一つは60歳前後から40歳前後は「共通語」と「方言」の使い分けを意識的に行うことが要求されかつその使い分けがコミュニケーション方略として意味をもつことを意識するライフステージにあるが、若い世代はまだそのような段階に至っていない、という解釈。もう一つは、若い世代は共通語化の行きついでいるため、「方言」と「共通語」の区別がつけにくくなっている、自身の用いることばを「共通語」であると認識しやすい、という解釈である。しかし、現時点では、これらの解釈を検証するに足る十分なデータをもっていない⁵。

最後に全体・近畿・東海ならびに首都圏とは異なるパターンを示した九州（図2d）・東北（図2e）における年齢による潜在クラスへの帰属確率の変動パターンについてみていく。

九州は「クラス1: 積極的方言話者」への帰属確率が基本的に高い地域であるが、60歳前後から40歳前後においては「クラス4: 積極的使い分け派」が優勢化する。40歳より若い年代ではふたたびクラス1が優勢化し、全体・近畿・東海・首都圏とは異なり、最若年層に向けてクラス1の帰属確率が高まる。

東北は各クラスへの帰属確率の差が小さく、クラス間の帰属確率の変動がやや複雑である。60歳前後より上の世代では「クラス2: 共通語話者」が「クラス1: 積極的方言話者」をわずかに上回るもの、60歳前後を境にクラス1への帰属確率が増加し、クラス2への帰属確率が減少する。40歳前後を変化点としてそれより若い世代に向けて若干クラス1への帰属確率が下がりクラス2の帰属確率が高まるが、クラス1の優勢は維持される。「クラス4: 積極的使い分け派」が40歳前後まではクラス2と並行的に増加するものの40歳前後をピークとして帰属確率を急激に下げ、最若年層ではクラス2を下回る。

九州と東北は上述のように異なる部分もあるが、年齢と各クラス帰属確率の変動をみていくと、大きくは「クラス4: 積極的使い分け派」から「クラス1: 積極的方言話者」に類型が移行しつ

⁵ ライフステージの観点から社会的活躍層がその前後の世代とは異なるふるまいをするという傾向は、山形県鶴岡市における調査データに基づく標準語化モデルである Nomoto-Egawa Model (江川 1973, 1977; 野元 1975; Yoneda 1997) のI期の形状からも確認される。社会的活躍層が前後の世代に先駆けて標準語化していく様子が示されている。I期においては世代効果だけでなく、教育程度と職業の効果が確認される点も、3.2節で述べた本稿の結果と共通する。標準語化と、積極的な方言使用というまったく異なる現象ではあるものの、社会的活躍層が先導的立場を示すという点、共通する。標準語化においては完成期(III期)の一段階前のII期では、若い世代ほど標準語化が進んでいるが、教育・職業の効果は確認されない。鶴岡データにおける標準語化のI期では性の効果（男性より女性が標準語得点が高い）も認められたが、本稿では3.2節で述べたように性の効果は認められなかった。

つあることがわかる。これは、将来これらの生育地群がクラス1の典型である近畿タイプに近づいていくことも想像させる。「方言」の価値が高まるととともに、少なくともイメージの明確な「方言」をもつと話者に意識される地域は、「積極的方言話者」の方向に傾いていく可能性を含意していると思われる。

ただし、明確なイメージの確立していない生育地方言や地域が帰属する確率の高い「クラス3: 消極的使い分け派」「クラス5: 判断逡巡派」については、年齢と各クラス帰属確率の変動について安定的な結果を用いた検討ができていない。よって、これらの類型がどのような方向性を示すのかについては、議論できない。

5. 今後の課題

2010年全国方言意識調査データに基づく話者の確率論的分類という手法を用いて、「方言」と「共通語」にかんする地域類型を抽出し、その背景ならびに先行研究で示された類型との比較を行なってきた。

無作為抽出による大規模データを用いた客観的な地域類型のこんにち版を話者の分類というミクロな視点から提案してきたことになるが、1.3.1節で述べたように、本分析で用いた約1,300というサンプル数は、「方言」という多様な現象を代表させるには多いとはいえない。とりわけ、地域差が大きな要因として効果をもつことは今回の分析からも明白に確認されており、分析結果の安定性を確保した地域ごとの対照を行なうためには、層別サンプル数の確保が重要となってくる。つまり、潜在クラス分析のような手法を用いた類型の抽出という観点からは、2010年全国方言意識調査データは、大規模調査であるとはいえ、細かい類型を得るにはややサイズ不足ということになる。世代効果や、時代効果を正確に考察するためには、さらに大規模な、同一項目の繰り返し調査というアプローチが必要であるが、これらについては今後の課題としたい。

参照文献

- Cochran, William G. (1977) *Sampling techniques*. Third edition. New York: John Wiley and Sons.
- Dayton, C. Mitchell and George. B. Macready (1988) Concomitant-variable latent-class models. *Journal of the American Statistical Association*. 83: 173–178.
- 江川清 (1973) 「最近二十年間の言語生活の変容」『言語生活』257: 56–63. 東京：筑摩書房.
- 江川清 (1977) 「第一章 地域社会の言語生活の変化・変遷」野元菊雄・野林正路監修『日本語と文化・社会2 ことばと社会』3–30. 東京：三省堂.
- 言語編集部 (1995) 「共通語との心理的距離」『月刊言語』24(12): 61. 東京：大修館書店.
- Hagenaars, Jacques. A. and Allan L. McCutcheon (2002) *Applied latent class analysis*. Cambridge: Cambridge University Press.
- 陣内正敬 (1996) 『地域語の生態シリーズ九州篇 地方中核都市方言の行方』東京：おうふう.
- 陣内正敬 (1999) 「次のような場面であなたが使うことばは？」佐藤和之・米田正人（編）135–141.
- 寿岳章子 (1978) 「標準語の問題」『岩波講座・日本語3 国語国字問題』52–55. 東京：岩波書店.
- 河西秀早子 (1981) 「標準語形の全国分布」『言語生活』354: 52–55. 東京：筑摩書房.
- McCutcheon, Allan L. (1987) *Latent class analysis*. Thousand Oaks, CA: Sage Publications.
- McLachlan, Geoffrey and David Peel (2000) *Finite mixture models*. New York: John Wiley and Sons.
- NHK放送文化研究所（編）(1997) 『現代の県民気質—全国県民意識調査一』東京：日本放送出版協会.
- 野林正路 (1971) 「少年・少女期の二重言語生活」『言語生活』239: 33–44. 東京：筑摩書房.

- 野元菊雄 (1975) 「第六章 年齢層と言語」波多野完治・野林正路 (編)『新・日本語講座 10 ことばと文化・社会』163–188. 京都: 汐文社。
- 岡太彬訓・守口剛 (2010) 『マーケティングのデータ分析—分析手法と適用事例一』東京: 朝倉書店。
- 佐藤和之 (1999) 「方言主流社会」「共通語中心社会」佐藤和之・米田正人 (編) 36–45.
- 佐藤和之・米田正人 (編) (1999) 『どうなる日本のことば』東京: 大修館書店。
- 多田治 (2008) 『沖縄イメージを旅する—柳田國男から移住ブームまで—』269. 東京: 中央公論新社。
- 田中ゆかり (2011) 『「方言コスプレ」の時代—ニセ関西弁から龍馬語まで—』東京: 岩波書店。
- 東條操 (1954) 「国語の方言区画」東條操 (編)『日本方言学』18–33. 京都: 吉川弘文館。
- 友定賢治 (1999) 「「つくられた」方言イメージと共に通語イメージ」佐藤和之・米田正人 (編) 97–100.
- Vermunt, Jeroen K. and Jay Magidson (2005) *Latent GOLD 4.0 user's guide*. Belmont, MA: Statistical Innovations Inc.
- Wand, Matt P. and M. Chris Jones (1997) *Kernel smoothing*. London: Chapman and Hall.
- Yoneda, Masato (1997) Survey of standardisation in Tsuruoka, Japan: Comparison of results from three surveys conducted at 20-year intervals. *Japanese Linguistics* (『日本語科学』) 2: 24–38. Tokyo: Kokusho Kankokai.

Regional Typology Based on Individual-level Clustering of Dialect Usage: A Latent Class Analysis of Nationwide Language Consciousness Survey Data

TANAKA Yukari^a

MAEDA Tadahiko^b

^aNihon University / Project Collaborator, National Institute for Japanese Language and Linguistics

^bThe Institute of Statistical Mathematics

Project Collaborator, National Institute for Japanese Language and Linguistics

Abstract

For the purpose of examining regional typology of dialect usage, we analyzed data from a nationwide survey of language consciousness targeting Japanese residents aged 16 and over. We adopted latent class analysis in order to derive the individual-level probabilistic clustering of language usage. Based on the results we tried to identify the regional correspondence to these clusterings (latent classes) and scrutinized the relationship between demographic characteristics of respondents and these classes. The analysis yielded five latent classes, which can be conceived as follows: "Class 1 Active Dialect Speaker", "Class 2 Common Language Speaker", "Class 3 Moderate Code Switcher", "Class 4 Active Code Switcher", and "Class 5 Indecisive Respondent". Among the demographic characteristics, hometown, occupation, educational background, and age all had a statistically significant effect on class membership (with effect size running from largest to smallest in this order), while gender and size of city of residence did not. The regional distribution of respondents' hometowns showed that certain classes are typical for respondents who grew up in specific regions: "Class 1 Kinki, Chugoku and Shikoku Area", "Class 2 Tokyo Metropolitan Area and Hokkaido", "Class 3 Northern Kanto, Koshin-etsu, Hokuriku, and Tokai Area", "Class 4 Okinawa, Kyushu, Tohoku and Chugoku Area", and "Class 5 Hokkaido". Occupation also showed a specific relationship to some of the classes. With regard to age, we observed a non-linear relationship both in the analysis of the total sample and in the analyses by regions. Generally speaking, each age-effect curve has two or more change points, most of which are observed between the ages of 35 and 40 and at around 60. Based on these results, we discuss the implications of regional typology and the relationship between the results of our study and previous studies.

Keywords: common language, regional dialect, language consciousness, regional typology, latent class analysis