

# 共同研究プロジェクト紹介 文字環境のモデル化 と社会言語科学への応用 言語変化は経年調査データから予測可能か?

著者	横山 詔一
雑誌名	国語研プロジェクトレビュー
号	6
ページ	27-37
発行年	2011-10
URL	<a href="http://doi.org/10.15084/00000683">http://doi.org/10.15084/00000683</a>

## 〈共同研究プロジェクト紹介〉 文字環境のモデル化と社会言語科学への応用

### 言語変化は経年調査データから予測可能か？

Quantitative Prediction of Language Change with Longitudinal Survey Data

横山 詔一 (YOKOYAMA Shoichi)

国立国語研究所 理論・構造研究系 教授

(Professor, Department of Linguistic Theory and Structure, NINJAL)

《要旨》 言語変化の経年調査データから将来の言語変化を数量的に予測するモデル（横山・真田 2010）について紹介した。このモデルは「臨界期記憶+調査年効果 → 共通語化」という図式にしたがって共通語化を説明・予測する。国立国語研究所が山形県鶴岡市を定点観測フィールドとして経年的に約 20 年間隔で過去 3 回実施した共通語化調査の大量データを、このモデルで解析した結果、アクセント共通語化などにおいて予測値と観測値が精度よく一致することが示された。

**Abstract:** This paper discusses the validity of a new statistical model of language change proposed by Yokoyama and Sanada (2010) in an analysis of data on dialect standardization in Tsuruoka City, Yamagata Prefecture. This model supposes that a regional dynamic change in language reflects a change over the years in linguistic memory traces which accumulate in the brains of the local residents during their life.

Employing this model, Yokoyama and Sanada analyzed longitudinal survey data from Tsuruoka City, collected three times at twenty years intervals, and statistically confirmed that the increasing standardization of dialect shows a close fit between observed data and model predictions.

## 1. はじめに

### 1.1 予測は学術レベルの試金石

「過去のデータから将来を数量的に正確に予測すること」、これは科学の目的の 1 つであろう。天文学者は、次の皆既日食が、いつ、どこで観測できるかをピタリと当てることができる。気象学者は、あした雨が降る確率をかなり正確に予測することができる。地震学者は、次の巨大地震が、いつ、どこで起きるか大ざっぱにしかに予測できない。では、言語学者は、いつ、どこで、どのような言語変化が生じるかを数量的に予測できるのだろうか？

たとえば、ある地域で長期間にわたって（科学的な方法で）収集された言語変化のデータを手にすることができれば、その地域における将来の言語変化を数量的に正確に予測できるかもしれない。そのようなデータは、はたして存在するのだろうか？ 答えは Yes である。山形県鶴岡市で約 40 年間にわたって収集された共通語化のデータが国立国語研究所に保存・蓄積されている。そのデータを解析すれば、鶴岡市の方言が 20 年後や 40 年後にどのくらい変化しているか（共通語化しているか）を数量的に予測できる可能性が開ける。以下、鶴岡データの説明と数量的予測の方法について解説しよう。

## 1.2 鶴岡データとは

国立国語研究所は山形県鶴岡市を定点観測フィールドとして「地域社会における方言の共通語化」に関する実態調査（以下、鶴岡調査という）を1950年（昭和25年）から約20年間隔で3回にわたって経年的に実施してきた。鶴岡市は、図1に示すように山形県の庄内地方南部に位置する。旧鶴岡藩（通称、庄内藩）の城下町で、文化・経済の中心都市として栄えた。時代小説の作家である藤沢周平の作品に登場する「海坂藩」のモデルと言われる。過去3回の調査時の人口は、第1回調査時（1950年）が約9万6千人、第2回調査時（1971年）が約9万5千人、第3回調査時（1991年）が約10万人と、大きな人口変動のない、日本に点在する中小都市の典型である。



図1 鶴岡市の位置

鶴岡調査は、第1回から第3回まで、いずれも住民基本台帳などにもとづいてランダム抽出されたサンプルをインフォーマント（調査対象者）としてきた。サンプリング調査は毎回約400名前後を対象としたほか、同一人物への追跡調査（パネル調査）もおこなってきた（国立国語研究所1953, 1974, 1994, 2007）。つまり、横断調査法と縦断調査法を組み合わせた調査デザインになっている。この調査方法は「コーホート系列法 (cohort sequential method)」と呼ばれ、生涯発達心理学、老年学、脳科学など幅広い分野で注目を集めている（横山・阿部2010, 横山・井上・阿部2010）。世界を見わたすと、海外で実施されたコーホート系列法による本格的な調査は米国の「シアトル調査」があるのみのものである。これは知能の生涯変化を探る目的で実施されている大規模プロジェクトで、米国シアトル市において1956年から7年ごとに経年的におこなわれている (Schaie 1996, Schaie and Hofer 2001, Schaie and Willis 2001)。鶴岡調査は1950年が第1回なので、シアトル調査よりも6年ほど早くスタートしており、世界最初のコーホート系列法による調査だと考えられる（横山2010）。このことから、鶴岡調査のデータが、量と質の両面で世界的に群を抜くレベルにあることが分かる。調査項目の基本部分は第1回から第3回まで同じものが使用されおり、41年間にわたる言語変化の実態をとらえることができる。そのデータは国立国語研究所報告書に一部が公開されている。

そのほか、国立国語研究所の研究情報資料センターは、国立国語研究所の理論・構造研究系が主導する共同研究プロジェクト「文字環境のモデル化と社会言語科学への応用」と密接に連携しながら、素データの一部（音声に関する31項目など）をデータベース化する作業を進め、学術研究に役立てている（阿部2010, 横山・阿部2010）。また、国立国語研究所のサイト <http://www2.ninjal.ac.jp/keinen/> では鶴岡調査だけではなく、岡崎調査など「ことばの定点経年観測」に関する資料が閲覧できる。

### 1.3 将来予測のリクツ

ここでは、どのようなリクツ（方法）によって将来予測ができるのかを横山・真田（2010）の研究を具体例として紹介する。この研究は4回目以降の調査が将来的に継続された場合の結果の一部を「生涯習得モデル」にもとづいて数量的に予測したものである。生涯習得モデルとは、簡単に言えば、「ある地域における言語変化の動態は、地域住民の脳内に生涯にわたって蓄積される言語記憶や言語運用能力の経年変化を反映している」という心理学的モデルである。

この理論枠を用いれば、過去に実施された調査のデータを統合して、新たな視座から共通語化の実態を概観することも可能になるかもしれない。また、脳内に蓄積された言語の知識・運用能力の時間的変容を追うという点では心内辞書 (mental dictionary) の生涯変化に関するモデルだとも言えよう。

## 2. 予測モデルの説明

記憶に関する最近の心理学的研究は、社会脳科学などと融合しながら、めざましい発展をとげている(横山・渡邊 2007)。横山・朝日・真田 (2008) は、記憶研究の知見を援用した言語変化予測モデルを作成し、国立国語

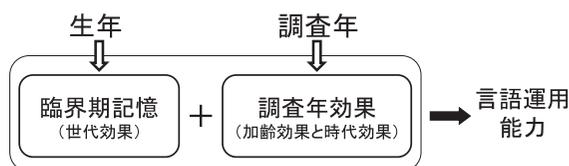


図2 生涯習得モデルの枠組み

研究所が愛知県岡崎市を定点観測フィールドとして実施した「敬語と敬語意識」の調査データを解析した。岡崎調査も鶴岡調査と同様に3回にわたって実施された。第1回は1953年、第2回は1972年、第3回は2008年であった。これは55年間にわたる定点経年調査である。岡崎調査データの解析に使われた横山・朝日・真田（2008）のモデルは横山・真田（2010）によって小さな改訂がくわえられ、図2に示すような「生涯習得モデル」として鶴岡市の共通語化予測に用いられた。生涯習得モデルの骨子は次の通りである。

- (1) 人間が言語習得期（10歳前後）に経験した言語接触の記憶は脳内に安定した痕跡を残し、終生保持されると考える。以下、これを「臨界期記憶」と呼ぶ。ある人が獲得した臨界期記憶は生年と深い関係がある。生年によって言語習得期にどの時代の社会を経験するのかが決まる。たとえば1960年代に生まれた人は、1970年代の経済高度成長時代に言語習得期を過ごしたことになる。その人が10歳前後に過ごした環境が、あるいは、その時代の言語の状況が臨界期記憶に影響を残すと考えれば、生年は臨界期記憶に相関する変数だと見なしてよい。

ちなみに、社会学や経済学のコーホート分析 (cohort analysis) では、世代の違いによる世代効果の変数として生年を扱う。言語変化を研究する場合は、臨界期記憶の影響によって世代効果が生じると仮定しても差し支えないと考えられるため、ここでは世

代効果と臨界期記憶の効果を区別しない。

- (2) 言語調査の結果には、「調査年効果」が含まれている。そして調査年効果は、時代効果と加齢効果を合算したものだと考えられる。時代効果とは、時代の違いによる影響を指す。鶴岡市における1950年調査と1971年調査では、テレビ普及率に格段の差があった。これは時代効果の典型例である。一方、加齢効果とは、年齢の影響を指す。1920年生まれの人が1950年に調査を受けたならば年齢は30歳であるが、同じ人が1971年に調査を受けると年齢は51歳になっている。

なお、コーホート分析の研究においては時代効果と加齢効果を分離する手法が提案されているが、「識別問題」と称される数理的に解決困難な問題があるため、ここでは時代効果と加齢効果を分離しないで調査年効果として扱う。

このように、生涯習得モデルは、言語習得が臨界期記憶と調査年効果によって支えられているとの仮定をおく。そして、言語運用能力は、臨界期記憶と調査年効果が合算されて発現すると考える（図2参照）。

言語運用能力は言語変化に関する現象と密接に関連しているので、図2の言語運用能力の項を共通語化に書き換えると、生涯習得モデルは式〔1〕のようになる。

$$\text{臨界期記憶} + \text{調査年効果} \rightarrow \text{共通語化} \quad [1]$$

先に述べたように、臨界期記憶は生年と、調査年効果は調査年と、それぞれ強い関係がある。よって、式〔1〕は式〔2〕のように書き換えてよい。

$$a_1 \times \text{生年} + a_2 \times \text{調査年} + b \rightarrow \text{共通語化率} \quad [2]$$

これは、国語と数学の2科目で学力テストをおこなった場合に、合計点をどのようにして求めるかという問題とよく似ている。たとえば国語の点数は2倍して数学の点数と足し算する、というような計算方法であれば、国語の得点の重みは2倍になる。式〔2〕は、生年と調査年という2つの変数に適当な重みをかけて足し算することにより、共通語化の程度が予測できるという言語変化予測式になっている。この予測式を実際に求める場合は、医学などでよく使われているロジスティック回帰分析という手法を用いる。ちなみに、ロジスティック回帰分析による言語変化予測は、文字環境のモデル化に関する研究のなかで基本的な理論枠が作られたという経緯がある（Yokoyama and Wada 2006, 横山・真田 2007）。

### 3. 言語変化予測の手順

鶴岡調査は、さまざまな調査を組み合わせることによって、共通語化現象を多角的に捉えてきた。その調査群の中で、過去3回のいずれでも実施されている中心的な調査が「共通語の調査」である。それらは、音声（音韻）、語彙、文法項目について、調査員が調査票を用いて個別面接方式で実施する調査であった。言語的事象以外では、性別・生年といった被調

査者の属性、マス・コミュニケーションとの接触度、方言や共通語に対する意識など、共通語化に関連する可能性のある種々の情報についても尋ねた。

横山・真田（2010）は、言語変化について後述の3つの典型例を記述することを第一の狙いとし、音声項目、アクセント項目などの区分はせず、音声項目 207「ネコ」、223「エントツ」、アクセント5項目、の3種の解析結果を例示した。

各項目の観点は以下の通りである。

- ・項目 207「ネコ」：非語頭におけるカ行有声化の有無（図3は第3回調査に用いた刺激図）
- ・項目 223「エントツ」：語頭の母音エにおける狭母音化の有無
- ・アクセント5項目「セナカ、ネコ、ハタ、カラス、ウチワ」：アクセントが共通語と同じであるか

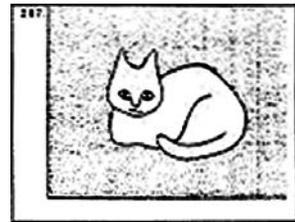


図3 「ネコ」項目の刺激図

### 3.1 生年と調査年について

先に示した式〔2〕の左辺は、生年と調査年の2つの説明変数を含む。生年は「11年刻み」と称される刻み幅とした（国立国語研究所 2007: 18）。調査年は実際に調査が実施された1950年、1971年、1991年の3回のほか、将来の変化を約20年間隔で予測するために2011年、2031年、2051年を用いた（後述の図4は除く）。

### 3.2 共通語化率について

式〔2〕の右辺にある共通語化率は次のようにして求めた。項目207と223については、共通語の音声で回答された百分率が国立国語研究所（2007）に掲出されているので、それらを共通語化率として用いた。アクセント5項目については、共通語アクセントの平均点が国立国語研究所（2007）に掲出されている。その平均点の数値を満点の5で割り算した後に100倍し、アクセント共通語化率とした。以下の予測式の解析に投入する目的変数は、これらの共通語化率を100で割り算して確率に変換したものであった。

### 3.3 予測式の推定

過去3回の共通語化率データを目的変数、生年と調査年を説明変数として、ロジスティック回帰分析をおこなった。生年と調査年の「重み」を最尤法で推定し、ステップワイズ法により5%水準で統計的に有意な説明変数を選択した。

表1の重みは回答に働く効果を示しており、数値が大きいほどその効果も大きい。正の場合は共通語化率を上昇、負の場合は下降させる働きがある。検定の結果、効果が有意でないと判定された重みについてはゼロ「0」と表示した。このように、研究者の主観で説明変数を選択するのではなく、統計的に有意な説明変数を選択する方法は、より客観的な基準にも

とづいた判断を可能にするものと考えられる。

表1 統計的に有意な重み

	$a_1$ : 生年の重み	$a_2$ : 調査年の重み	$b$
項目 207 「ネコ」	+0.0475892	0	-90.0756503
アクセント 5 項目	+0.0254825	+0.0264587	-102.4427798
項目 223 「エントツ」	+0.0779051	-0.0501265	-49.3547537

#### 4. 言語変化予測の結果

言語変化をグラフの形状から読み取るにあたり、横軸をインフォーマントの生年、縦軸を共通語化率とする図を描いた。それらを視察した結果、後で説明するように、鶴岡市の共通語化は以下の4つのタイプに分かれて進展しつつあると考えられる。なお、これらのタイプのいくつかは、先行研究においてすでにその存在が指摘されているほか、言語変化がS字カーブを描くことを示した論考も存在する（江川 1973, Aitchison 1991, 井上 2000, 井上・江川・佐藤・米田 2009）。ここでも先行研究にならって言語変化のS字カーブを念頭におきながら、将来の変化をさらに高い精度で予測できるようにする新たな方法を検討した。

##### 4.1 生年だけが共通語化を促進させるタイプ

図4に音声項目207「ネコ：非語頭におけるカ行有声化の有無」の結果を示す。表1にあるように、この項目は調査年の効果が統計的に有意ではなかった。生年の効果は統計的に有意なので、生年だけを説明変数とする予測式から得た予測値が観測値とうまく一致している。このタイプは、後の調査でも生年が同じ世代の回答は変わらないので、共通語化のカーブは一本につながる。つまり、「言語習得期に獲得した言語記憶痕跡は残りの生涯にわたってほぼ変わらずに保持される」ことが経年調査によって実証された典型例である。

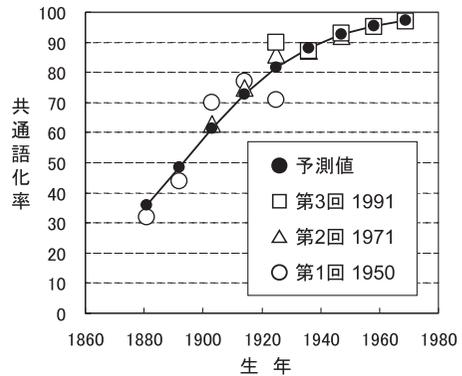


図4 「ネコ」項目207の共通語化予測

予測値と観測値のズレ（誤差）がもっとも大きかったのは1950年に調査を受けた1925～1935年生まれの子孫で、誤差の大きさは10%であった。そのほかのほとんどの群は誤差が5%以内にとどまった。

この項目の共通語化率は天井効果の様相を呈しており、1991年の時点ですべての世代で共通語化はほぼ完了（収束）したように見える。2011年、2031年、2051年の予測値はいずれの世代もほぼ100%になると予想されるため、図4では、それらの数値を割愛した。

なお、「ネコ」をアクセントの面から検討すると、図4とは違う共通語化の進行状況を描く様相が読み取れる。その点については、次のアクセントに関する5項目の結果のところでも詳しく述べる。

#### 4.2 生年と調査年の両者が共通語化を促進させるタイプ

図5に示す共通語化率はアクセントに関する5項目から算出されたものであり、単一項目による数値よりも信頼性が高い可能性がある。なぜなら、測定誤差が項目間で相殺されて真値に近い数値になっていると考えられるからである。

グラフ全体をながめると、第3回までは共通語化の進行速度が次第に加速していったようすが見て取れる。表1に示されているように、生年の効果と調査年の効果の両者とも統計的に有意であった。そのため、生年のほか調査年も説明変数に加えた予測

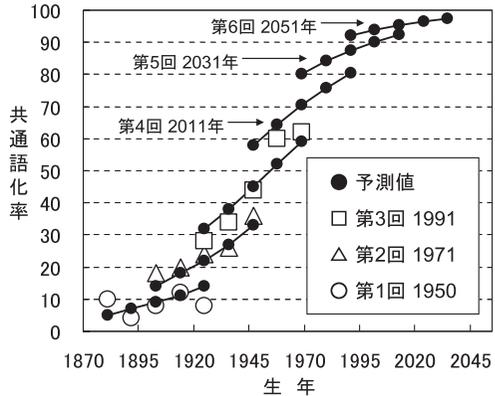


図5 アクセントの共通語化予測

式から得た予測値が観測値とうまく一致し、共通語化のカーブは一本にはつながらない。このタイプは、たとえ生年が同じ世代であっても、後の調査になればなるほど共通語化率は高くなる。これを「跳ね上がり現象」という。すなわち、同じ話者でも年齢が上になるにつれて共通語化していくことを示している。

予測値と観測値のズレがいちばん大きかったのは1991年に調査を受けた1958～1968年生まれの群で、誤差の大きさは8%であった。そのほかの大部分の群は誤差が5%以内にとどまった。あわせて、2011年8月現在でまだ調査が実施されていない4回目(2011年)から6回目(2051年)までの予測結果も図5には示してある。

この結果は、「見かけ上の時間 (apparent time) による調査」は変化の速度を過小評価する傾向があるという説 (Sankoff 2006) を強く支持するものである。見かけ上の時間とは、「臨界期に習得した言語運用能力が生涯にわたる言語運用を決定づける」という仮定のもと、若年層から老年層までのいろいろな年代層 (世代) を対象に1回の調査を実施し、そこから過去の言語使用の姿を推定しようとする研究手法を指す (Labov 1972, Chambers 1998)。見かけ上の時間による調査は、ある時点で1回だけしかデータを収集しないので、たとえば図5に示す第2回調査時点の1971年に1回だけ調査をおこなったとすると、跳ね上がり現象を観測できない。そのため、変化の速さを過小評価してしまうことになる。

各項目を個別に検討した結果、203a「セナカ：共通語 LHH, 方言 LHL」、207a「ネコ：共通語 HL, 方言 LH」、210a「ハタ：共通語 LH, 方言 HL」、228a「カラス：共通語 HLL, 方言 LHL」、231a「ウチワ：共通語 LHL, 方言 LLH」のいずれのグラフもこのパター

ンに属し、生年と調査年の効果が統計的に有意であった。207a「ネコ」は、非語頭におけるカ行有声化の有無（図4）とは違うパターンになっている。つまり、音韻の共通語化とアクセントのそれは、変化のパターンが異なっている場合があるようである。

#### 4.3 生年が促進効果で調査年が抑制効果を示すタイプ

音声項目 223「エントツ：語頭の母音エ」における狭母音化の有無の結果を図6に示す。分析の結果、生年のほか調査年も説明変数に加えた予測式から得た予測値が観測値とうまく一致した。この項目で予測値と観測値のズレがいちばん大きかったのは1971年に調査を受けた1914～1924年生まれで、誤差の大きさは7%であった。そのほかのほとんどの群は誤差が5%以内にとどまった。

ただし、ここで注目すべきは調査年の重み  $a_2$  が負の値になっていることである（表

1参照）。これは調査年が新しいほど共通語化とは逆の方向に進んでいることを意味している。このタイプは、たとえ生年が同じ世代であっても、後の調査になればなるほど共通語化率は低くなる。これを「引き戻し現象」という。すなわち、同じ話者でも年齢が上になるにつれて方言化することを示しており、先にふれた「見かけ上の時間による調査は変化の速度を過小評価する傾向がある」という説(Sankoff 2006)とは逆になった。たとえば第2回調査時点の1971年に1回だけ調査をおこなったとすると、引き戻し現象を観測できないので、変化の速さを過大評価することになる。

#### 4.4 その他

そのほかの項目のグラフをながめると、第2回調査時の共通語化率がピークで、第3回調査時の共通語化率は第1回調査時の付近まで低下するものがある。これは、調査年によって波動的に周期変動しているのかもしれない。このタイプのデータ解析は式〔2〕の右辺を2次式や3次式にすることで対応できる。ただし、複雑系科学のカオス理論などを視野に入れる必要もあると考えられるため、別の機会に検討を譲る。

### 5. まとめ

ここでは、経年的に実施された共通語化調査データのなかに潜在している生年と調査年の効果を分離・抽出し、調査年が（生年とは逆に）方言化に寄与している実例などを統計的根拠とともに示した。図2のようなモデルを想定することで、言語変化を経年調査データから

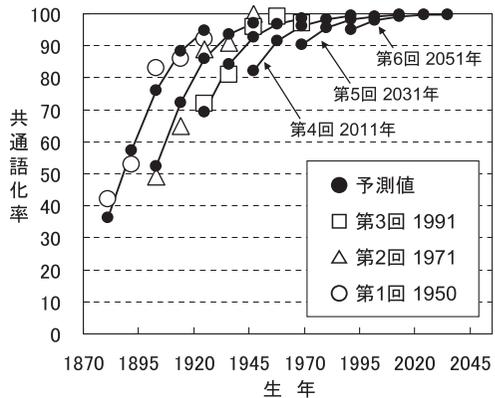


図6 「エントツ」項目 223 の共通語化予測

数量的に予測できる場合もあるようだ。

朝日・松田（2008）は、一般に過去の経年調査（実時間調査）では言語変化の抽象的モデルの提示とパーセンテージの吟味にとどまっていると指摘している。それは共通語化の経年調査においても例外ではなく、説明変数（独立変数）として生年と調査年、あるいは年齢層と調査年を取り上げ、目的変数（従属変数）の共通語化率が描く軌跡をグラフで視察する段階を脱していないところもある。そこから一歩踏み出し、生涯習得モデルの視点から言語変化予測式を用いた解析をおこなえば、ある地域社会の住民が当該の調査時期に脳内に蓄積していた言語記憶の平均像を読み取ることが可能になるであろう。すなわち、鶴岡市で観測された共通語化の経年変化データから、鶴岡市民の脳内に蓄積された共通語の記憶痕跡の動態を推測することも難しくはないだろう。言語変化の背後にある潜在構造をあぶり出すには上記のような理論が必要である。

2011年冬ごろには、統計数理研究所の調査科学研究センターが主体となって国立国語研究所と共同で鶴岡調査の4回目のランダムサンプリング調査とパネル調査を実施する予定である。その調査が実現すれば、先に示した予測のグラフがどのくらいの的中しているかを検証できる。そもそも、鶴岡調査は、1950年のスタート時点から統計数理研究所と密接に連携しながら共同研究を進めてきたという特色をもつ。統計数理研究所との共同研究体制が確立していたからこそ、当時の日本で黎明期にあったランダムサンプリング調査の実施が可能になったという事実を忘れてはならない。

さらに、鶴岡調査の未来に眼をむけると、過去の調査とは違う新たな視点に立ってテーマを設定するべきだと考える。国立国語研究所の共同研究プロジェクト「文字環境のモデル化と社会言語科学への応用」には、鶴岡調査の将来像を検討する研究チームがある。その研究メンバー（井上史雄、佐藤亮一、佐藤和之、米田正人、阿部貴人、横山詔一ほか）は、鶴岡調査を日本語研究のなかで今後50年間は継続できることを目指して過去3回の調査の問題点を洗い出し、解決策を議論してきた。そこでは、「共通語がほぼ全国に行き渡った現在においては、共通語化そのものではなく、共通語と方言の使い分け行動の実態、およびそのような行動をとらせている言語使用者たちの意識を科学的に知る方法を追究することに調査の目的をシフトさせるのがよい」という共通理解が生まれている。この点に関する共同プロジェクトを、いかに企画し、実行するかを国立国語研究所全体で検討する時期が到来したように思える。

## 【謝 辞】

ここで紹介した内容は、国立国語研究所で2009年10月にスタートした共同研究プロジェクト「文字環境のモデル化と社会言語科学への応用」の成果の一部である。プロジェクトのメンバーに感謝の意を表する。また、本稿の詳細は横山詔一・真田治子（2010）に掲載されている。真田治子氏（立正大学経済学部教授）にも感謝の意を表する。

## 参考文献

- 阿部貴人 (2010) 『鶴岡調査回答データ説明書トライアル版 ver.0.1』国立国語研究所 (理論・構造研究系共同研究プロジェクトと研究情報資料センターの連携研究内部資料).
- Aitchison, Jean (1991) *Language change: Progress or decay?* 2nd edition. Cambridge: Cambridge University Press.
- 朝日祥之・松田謙次郎 (2008) 「敬語研究と実時間的言語変化研究との接点を求めて」『社会言語科学』11: 39-50.
- Chambers, Jack (1998) Social embedding of changes in progress. *Journal of English Linguistics* 26: 5-36.
- 江川清 (1973) 「最近二十年間の言語生活の変容—鶴岡市における共通語化について—」『言語生活』257: 56-63.
- 井上史雄 (2000) 『東北方言の変遷』東京: 秋山書店.
- 井上史雄・江川清・佐藤亮一・米田正人 (2009) 「音韻共通語化の S 字カーブ—鶴岡・山添 6 回の調査から—」『計量国語学』26: 269-289.
- 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』, 国立国語研究所報告 5. 東京: 秀英出版.
- 国立国語研究所 (1974) 『地域社会の言語生活: 鶴岡における 20 年前との比較』, 国立国語研究所報告 52. 東京: 秀英出版.
- 国立国語研究所 (1994) 『鶴岡方言の記述的研究: 第 3 次鶴岡調査報告 1』, 国立国語研究所報告 109-1. 東京: 秀英出版.
- 国立国語研究所 (2007) 『地域社会の言語生活: 鶴岡における 20 年間隔 3 回の継続調査』国立国語研究所.
- Labov, William (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Sankoff, Gillian (2006) Age: Apparent time and real time. In: Keith Brown et al. (eds.) *Encyclopedia of language & linguistics*, 2nd edition, 110-116. Oxford: Elsevier.
- Schaie, Warner (1996) *Intellectual development in adulthood: The Seattle longitudinal study*. New York: Cambridge University Press.
- Schaie, Warner and Scott Hofer (2001) Longitudinal studies in aging research. In: James E. Birren and Warner Schaie (eds.) *Handbook of the psychology of aging*, 5th edition, 53-77. California: Academic Press.
- Schaie, Warner and Sherry Willis (2001) *Adult development and aging*, 5th edition. New Jersey: Prentice Hall. [シャイア & ウィルス, 岡林秀樹 (訳) (2006) 『成人発達とエイジング 第 5 版』東京: プレエン出版.]
- 横山詔一 (2010) 「音声共通語化の予測と検証」『日本音声学会第 24 回全国大会予稿集』3-11.
- 横山詔一・阿部貴人 (2010) 「地域社会の共通語化に関する多変量解析モデル」『情報処理学会研究報告: 人文科学とコンピュータ研究会報告』2010(4): 33-48.
- 横山詔一・朝日祥之・真田治子 (2008) 「記憶モデルによる敬語意識の変化予測」『社会言語科学』11: 64-75.
- 横山詔一・井上文子・阿部貴人 (2010) 「言語生活の生涯変化は多変量解析で予測可能か: アクセントや敬語意識を例に (企画委員セッション「言語変化を追跡・予測する: 言語の定点経年調査から」の一部)」『日本行動計量学会第 38 回大会抄録』70-73.
- 横山詔一・真田治子 (2007) 「多変量 S 字カーブによる言語変化の解析—仮想方言データのシミュレーション—」『計量国語学』26: 79-93.
- 横山詔一・真田治子 (2010) 「言語の生涯習得モデルによる共通語化予測」『日本語の研究』6(2): 31-45.
- Yokoyama, Shoichi and Yukiko Wada (2006) A logistic regression model of variant preference in Japanese kanji: An integration of mere exposure effect and the generalized matching law. *Glottometrics* 12: 63-74.
- 横山詔一・渡邊正孝 (2007) 『記憶・思考・脳』東京: 新曜社.

## 横山 詔一 (よこやま・しょういち)

国立国語研究所理論・構造研究系教授。博士 (心理学) (筑波大学)。上越教育大学助手, 国立国語研究所領域長, 同研究所グループ長を経て, 2009 年 10 月より現職。

主な著書・論文: 「表記と記憶」(心理学モノグラフ No.26, 日本心理学会, 1997), 『記憶・思考・脳』(共著, 新曜社, 2007), 「言語の生涯習得モデルによる共通語化予測」(共著, 『日本語の研究』6(2), 2010), 「記憶モデルによる敬語意識の変化予測」(共著, 『社会言語科学』11, 2008), A logistic regression model of variant preference in Japanese kanji: An integration of mere exposure effect and the generalized matching law. (with Yukiko Wada, *Glottometrics* 12, 2006).

受賞：日本教育工学会論文賞（日本教育工学会，1997），徳川宗賢賞（優秀賞）（社会言語科学会，2010）。  
社会活動：社会言語科学会理事，社会言語科学会研究大会発表賞選考委員会委員長，計量国語学会理事，日本心理学会認定心理士認定委員会副委員長。

### 基幹型共同研究プロジェクト「文字環境のモデル化と社会言語科学への応用」

プロジェクトリーダー 横山詔一（国立国語研究所 理論・構造研究系 教授）

#### プロジェクトの概要

パソコンや携帯電話で文字を打つとき，私たちは「桧—檜—ヒノキーひのき」のような変換候補から1つの表記を瞬間的に選択している。では，どの文字表記が選択されやすいのかを，確率理論などで数量的に予測することは可能なのだろうか。このような問題意識のもと，日本語の文字表記について，日本語学の知見に加えて，認知科学，計量経済学などの優れた発想・理論も参照しながら研究を展開し，文字環境（文字レキシコンを含む）の質的・量的モデルを作成する。

また，この文字環境モデルは，敬語の経年変化や，地域社会で進行中の共通語化などの研究にも応用できる。とりわけ，愛知県岡崎市の敬語研究や，山形県鶴岡市の共通語化研究については，研究情報資料センターや言語変異研究系と連携しながら検討をおこない，言語変化理論の統計的検証を目指したデータ整備を進めて，社会言語科学や計量言語学の発展に寄与する。その成果は，言語習得研究のほか，脳科学や老年学研究にも新たな地平を拓く可能性があり，既存の分野の枠を超えた学際領域の創出につながるものと期待される。このように，本研究は学問領域の創成又は既存の分野の枠に収まらない学際領域の融合を目指している。