

埼玉学園大学・川口短期大学 機関リポジトリ

言語変化のS字カーブ：解析手法の比較とその適用事例

著者	真田 治子
雑誌名	埼玉学園大学紀要. 人間学部篇
巻	8
ページ	1-11
発行年	2008-12-01
URL	http://id.nii.ac.jp/1354/00000775/



言語変化の S 字カーブ

—— 解析手法の比較とその適用事例

S-curve Model of Language Change:
Methods and Examples

真田 治子
SANADA, Haruko

Language change is represented on a graph as a “slow-quick-quick-slow” S-curve. Fitting an S-curve model, language change data can be expected to summarize and analyse past trends and to predict future trends. This study discusses Piotrowski-Altman’s model and a model using a logistic curve, which are the major ones employed in the limited number of S-curve studies. We also discuss cases — a study of honorifics in Okazaki city, the standardization of modern Japanese in Tsuruoka city, and Chinese character words spreading in 19th and 20th century Japanese — which employ a multiple logistic regression with a multifactor approach. The requirements of an S-curve model for a study of language change are considered.

1. 各種言語調査にみられる言語変化の S 字カーブ

新しい語が人々の間に浸透していくときに、初期の段階ではじわじわと、そしてある段階で急激に広まり、それが一定の範囲に達するとまた伝播の速度が落ちていく、という現象はしばしば観察される。新語の流行だけでなく一般に言語変化の進行では、その浸透の比率が時間の経過と共に S 字カーブの軌跡を描いて上昇していくと考えられている。Aitchison は、これをダンスのステップになぞらえて「slow, quick, quick, slow」と表現している(1991)。このほか Denison も S 字カーブのモ

デルについて考察を行っている(2003)。

例えば図 1 は、英語の動詞「sneak」の過去形が「sneaked」から「snuck」に交替する様相を、カナダ Golden Horseshoe 地域で話者の年代と合わせて調査したものである (Chambers, 2006)。縦軸は「snuck」を使う人の百分率、横軸は話者の年代を生年順に 10 年ごとにまとめている。年代が若くなるにつれて「snuck」を使う人の割合が高くなるだけでなく、その浸透の速度は緩急があり全体として S 字カーブになることを示している。横軸の話者の年代を 1920 年代生まれ、1930 年代生まれ、……と左から生年順にすることで、調査参加者の世代からその生年が客観的な時系列データと

キーワード：言語変化、S 字カーブ、多変量ロジスティック回帰分析、Piotrowski-Altman カーブ、哲学字彙
Key words：Language Change, S-curve, Multiple Logistic Regression, Piotrowski-Altman’s curve, Tetsugaku
Jii

して算出でき、1回の共時的調査から過去の言語使用の姿をいわば通時的に得られる。このような「見かけ上の時間（apparent time）」から語形交替がS字カーブを描いて進行していることがおおそ観察できる。

言語の伝播だけでなく、人口の増加過程、水産資源の増殖過程、ファッションの普及過程などもS字カーブを描き、図2のようなロジスティック曲線（logistic curve）とよばれる関数がよく当てはまるといわれている。

日本語では、国立国語研究所がすでに60年以上、定期的に愛知県岡崎市、山形県鶴岡市などで定点観測的に言語データを収集している。井上史雄（2000）はこの鶴岡市のデータに自身による調査データを加えた40年以上の実時間調査によって「共通語化の過程」という言語変化のS字カーブを観測し、記述的研究をしている。

しかし言語変化の記述的研究ではフィールド調査、文献調査などで多くの時系列データが採取されているのに対して、言語変化のS字カーブを計量的に分析する手法はまだ非常に少ない。また「見かけ上の時間」の考え方では、「若年で習得した言語は終生変化しない」ことが前提となっており、年代が上がる

ことでの社会的役割の変化や時代による影響を1回の調査結果からだけで考察することはなかなか難しい。さらに細かな分析のためには、経年調査の結果を忍耐強く積み重ねる必要がある、同時にそれらを記述するだけでなく集積したデータを複眼的に分析する計量的な手法の開発も重要である。

言語学において、言語データの収集と記述的研究は最も基礎的な手法で、その性質から帰納的分析を行うのが一般的である。言語データを回帰分析にかけるということは、収集したデータの傾向を集約しデータを読み解きやすくする狙いがあるが、過去の傾向の延長として予測を行うこともできる。言語学における予測は、今後の調査に向けて確実に一定の結果が得られる箇所をピンポイントで絞り込むためにある。方言のようなフィールド調査であれば次にいつどのくらいの規模で誰を調査すればよいか、文献調査であれば次は小説や雑誌などどんな分野の資料からデータを集めればよいかの重要な判断材料になる。これは次の言語調査に向けて仮説を立てる行為であり、演繹的分析の始まりとなる。その意味で言語データの回帰分析と精度の高い予測は、帰納的分析から演繹的分析に切り替え

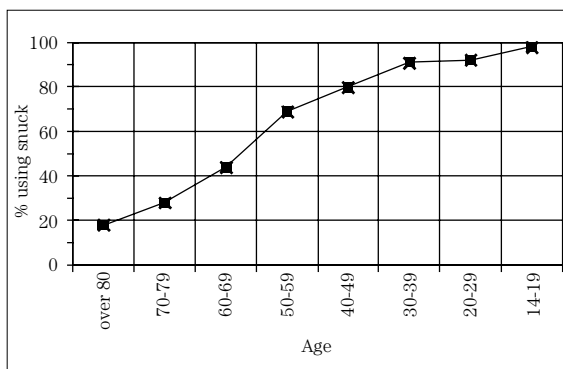


図1 カナダ英語における語形交替の例 (Chambers 2006)

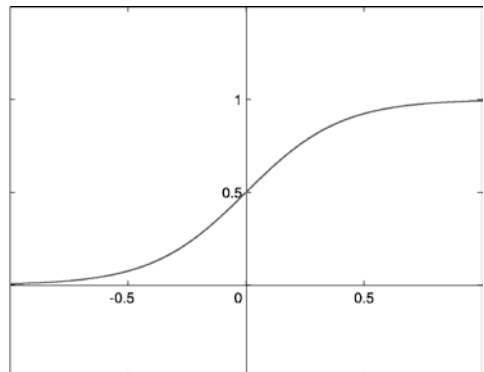


図2 ロジスティック曲線の例 (縦軸上限値1.0、下限値0.0)

て研究をさらに進展させていく支援ツールであるといえる。

このような背景をふまえて本稿では、言語変化のS字カーブの分析手法についての理論を比較検討する。さらに複数の言語データを重ね合わせて分析する多変量ロジスティック回帰分析の特徴を、文献から収集した学術用語のデータを用いて検討したい。

2. 言語変化のS字カーブを分析するためのモデルと解法

言語変化のS字カーブに関する理論研究では、ロシアを含む欧州諸国においてAltmann, Buttlar, Rott & Strauss (1983) の近似カーブがよく知られている。Altmann学派の方法は言語変化の動的活力 (dynamism) を、ある状態から次の状態へ移行するものととらえ、その変化をもともと微分方程式の形で表現しようとしたものである。この考え方は p を「比率」、 t を「時間」、 A 、 B 、 C 、 K を定数としてモデル①～③の式で表現されている。

Altmannのモデルは上限値を設定でき、増

モデル① 上限値を1、下限値を0とした単純増加のPiotrowski-Altmann S字カーブ

$$p = 1 / [1 + A * \exp(-K * t)] \quad (\text{式1})$$

モデル② 上限値(c)を予め100%以下に設定した単純増加のPiotrowski-Altmann S字カーブ

$$p = C / [1 + A * \exp(-K * t)] \quad (\text{式2})$$

モデル③ 一定の値まで増加したあと、減少に転じるPiotrowski-Altmann S字カーブ

$$p = 1 / [1 + A * \exp(-K * t + B * t^2)] \quad (\text{式3})$$

加・減少を繰り返す循環型・波動型のデータにも対応している。このうち式1のカーブは変形させると、式4に示す最も単純な形式のロジスティック曲線 (モデル④) の特別な形になることがわかっている (横山・真田

モデル④ 2変数のロジスティック曲線

$$p = 1 / [1 + \exp(-(K * t + a))] \quad (\text{式4})$$

$$\log\{p / (1-p)\} = K * t + a \quad (\text{式5})$$

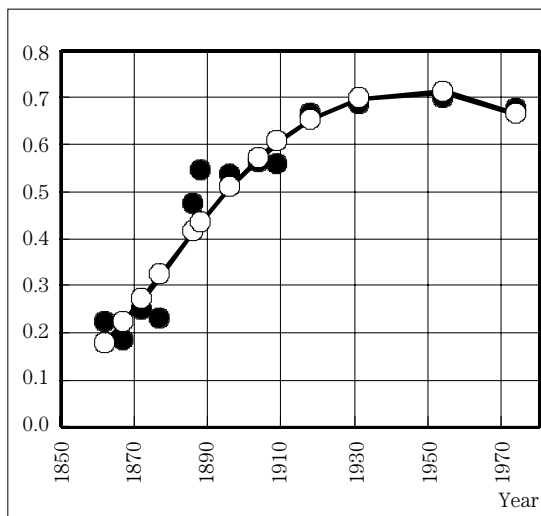


図3 明治期学術漢語の辞典における消長 (真田 2002) (モデル③による)

2007a)^(注1)。ロジスティック曲線は基本的に単調増加のS字カーブを示す曲線である。式中 p は「比率」、 t は「時間」、 a 、 K は定数を示す。式4は、対数の底を e とした式5の形式で表現されることもある。

これらのモデルは、『哲学字彙』（井上・和田垣・国府寺・有賀1881）収載の学術漢語の外国語辞書における消長（真田2002、図3）、新聞における外来語の増加過程（橋本2006）、文字論の研究（Yokoyama & Wada 2006）などの回帰分析に適用されて、言語変化の進行状況の分析と予測に成果をあげている^(注2)。

さらにロジスティック曲線では説明変数として時間だけでなく質的変数も含めた複数の変数を扱うことができる。これは多変量ロジスティック回帰分析とよばれる手法で、次節で実際の言語のデータを用いた事例について述べる。ロジスティック曲線の一般的な形(モデル⑤)を式6・式7に示す。 Z は $a1*X1 + a2*X2 + a3*X3 + .. b$ のような線型結合(重回帰式の形)とおく。多変量ロジスティック

<p>モデル⑤ (多変量) ロジスティック曲線</p> $p = 1 / [1 + \exp(-Z)] \quad \text{(式6)}$ $\log\{p / (1-p)\} = Z \quad \text{(式7)}$

曲線の場合は Z にあてはめる重回帰式に、2つ以上の変数を用いる。また年齢などの数値データだけでなく場面差などの質的データを変数に設定することもできる。この式6および式7の Z を $a*t+b$ とおいたものが上述の式4・式5で、あてはめる変数は時間(t)1つだけになっていた。つまり式6・式7の変数の数を限定した形が式4・式5といえる。

またロジスティック曲線にはモデル⑥のように下限値と上限値をデータから推定するモデルもある。

<p>モデル⑥ 上限値 $c1$・下限値 $c2$を推定するロジスティック曲線</p> $p = c1 - \{(1-c1-c2) / [1 + \exp(-(K*t+a))]\}$ <p style="text-align: right;">(式8)</p>

図4はモデル①からモデル④とモデル⑥を模式的に示したものである。実際の曲線の立ち上がり度合いは係数によって異なる。またモデル②とモデル③の飽和点、モデル⑥の上限値と下限値も係数によって異なる。

多変量ロジスティック曲線(モデル⑤)は医学、心理学の分野などでは広く用いられてきたが、前節で上述した通り、言語学での適用事例はほとんどない。そこでここでは、言

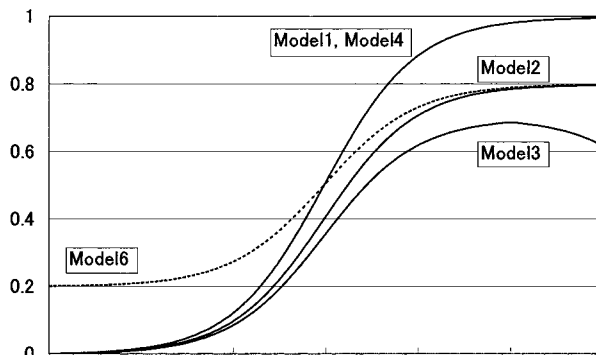


図4 モデル①からモデル④およびモデル⑥の模式図

語学ではどのような研究に適用できるかを考察したい。

言語変化の多変量S字カーブの分析では、言語変化が一定の時間で推移すると考えられるもので、その推移の様相が具体的なデータとして採取できるものなら適用できる。推移を示す中心的なデータの他に、周辺の変数として加味することで、より詳細な分析や複合的な分析が可能になるだけでなく、従来S字カーブととらえられていなかった事象に新しい分析の視点が期待できる。例えば、方言の被調査者の世代と学歴、あるいは経年調査の場合は被調査者の世代と調査年、また時系列の文献調査の場合は文献の書かれた時代とジャンルなどが変数に設定できよう。

3. 多変量ロジスティック回帰分析を使った言語研究—フィールド調査と文献調査

多変量ロジスティック回帰分析は最近特にフィールド調査の分析で成果をあげている。

国立国語研究所の愛知県岡崎市における敬語調査では、過去2回の大規模調査に、範囲

を限定したウェブ調査の結果を加えて非常に高い精度で傾向を分析した（横山・朝日・真田2008^(註3)）。この事例では、被調査者の生年を $X1$ 、3回の調査年を $X2$ とし、 Z を $a1*X1 + a2*X2 + b$ としてモデル⑤にあてはめている。ウェブ調査は3回目の大規模調査のための実験的な予備調査であったが、限られたデータ量で回帰曲線への高い適合性を示したため、この分析結果は2007年度から2009年度にかけて行われる3回目の本調査のための重要な資料となった。

同じく国立国語研究所の山形県鶴岡市における共通語化調査でも、過去3回の調査の調査結果が多変量ロジスティック回帰分析にかけられている（横山・真田2008）。この事例でも、被調査者の生年を $X1$ 、3回の調査年を $X2$ とし、 Z を $a1*X1 + a2*X2 + b$ としてモデル⑤にあてはめている。この分析では、第3次調査報告書（2007）において「加齢により共通語化率が低下するという奇妙な現象」と記されていた現象が、実はS字カーブの一部をなしており、それによって同じ話者でも年齢が上になるにつれて方言化が進行している

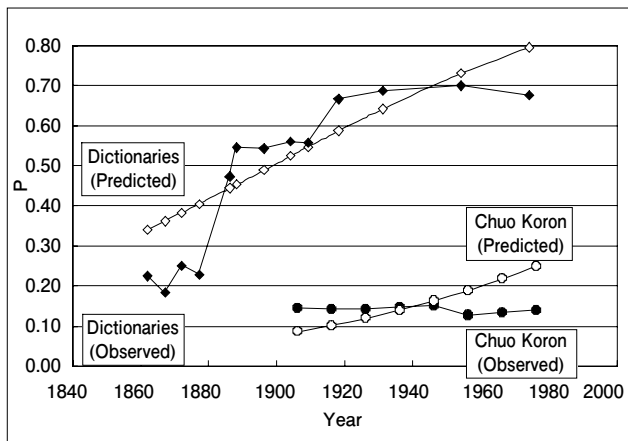


図5 明治期学術漢語の辞典と雑誌における消長（モデル⑤による）

ことを証明した。

ここでは、文献から採取した時系列の言語データに多変量ロジスティック回帰分析がどのように適用できるか、その可能性と問題点を検討する。

図5は明治期に広まった学術漢語が辞典と雑誌『中央公論』に採用される様子を示したもので、図3の辞典のデータに雑誌『中央公論』のデータを追加して2か所に散ったデータを一つのモデル（式）で統一的に再分析したものである。黒い丸が実データ、白い丸が理論値を示している。

辞典と雑誌コーパスでは言語資料としての性質が異なるが、このような質的変数では例えば辞書を0、雑誌を1と区別したダミー変数を用いることによって多変量ロジスティック回帰分析にかけることができる。そこで資料の出版年を $X1$ 、資料の種別を $X2$ として辞典と雑誌における漢語約2,000語の浸透率を p とし、モデル⑤にあてはめて以下の式を得た。データの曲線へのあてはまり度を示す R^2 は0.859であった。

$$p=1/(1+\exp(-0.018*X1+2.495*X2+34.169))$$

言語調査における質的データは、ほかには話者の性別、学歴、あるいは採取した言語資料の書き手の性別などが考えられる。また複数の変数を用いる解析では重回帰分析が知られているが、上述のように多変量ロジスティック回帰分析では上限値が100%、下限値が0%と固定されているため、その範囲からはみ出すことなく解析できる。

一方 $R^2=0.859$ というのは悪い値ではないが、図5をみると実データの変化と曲線とはやや乖離しているところもみられる。モデル④・モデル⑤に示したロジスティック回帰曲線は

$p=1$ に向かって単調増加することを前提にしているため、辞書のデータが戦後に入ってやや減少傾向にあることが反映されない。辞書は古語や規範的な形を示す役割もあるが、その時代に一般に通用している語を広く掲出する資料であるため、コーパスなどの資料に比べると異なり語数は多くなる。一方雑誌コーパスのような「語の使用」を記録した資料ではその時代のすべての語を掲出することは現実的には考えにくく、 $p=1$ を想定すること自体に無理がある。また循環型・波動型に対応しているモデル3は2変数を前提にしているので、辞書と雑誌と浸透率の3変数を扱うことはできない。

このような場合は最も単純に回帰直線を使うという考え方もあるが、言語は長い時間の中で増加と減少、定着と消失を繰り返す性質があるので、現在の傾向を把握し予測を視野に入れる場合には単調増加のモデルでは不十分で、言語学に適した曲線のモデルを考えるべきであろう。上限値と下限値を設定したモデルや循環型・波動型のモデルへの展開が必要である。また既存のS字カーブのモデルは前半の助走と後半の飛躍の形状が対称になっているが、実際の言語変化では、非常に長い時間助走段階にあり、ある時突然飛躍するものもある（Aitchison 2000）。カーブの形状のどこか一部分がデータに合うというだけでなく、将来的にもっと自由度の高いモデルに展開できることが望ましい。

しかし言語のS字カーブの計量的研究や理論的研究はまだ非常に少ないため、現在私たちが適用できるS字カーブのモデルは限られている。そこでそれらのモデルを資料の性格や研究の目的に合わせて使うために、モデルの特徴を次節で比較検討する。

4. Piotrowski-AltmanのS字カーブとロジスティック曲線の比較・検討

2節で述べたAltman学派のS字カーブのモデル(①②③)とロジスティック曲線(④⑤⑥)には各々特徴がある。差異は主に実務的な解法の点と理論の背景の思想的な点の二つにまとめられる。表1には6つのモデルの比較を示した。

・解法に関する問題

Altman学派のS字カーブのモデルは最小二乗法による推定を前提としており、解法を理解していれば線形回帰に変形させて解くことができる。多くの場合言語のデータは自然科学などと違い、非常に高い精度を必要としていないので、多少の誤差を考慮すれば特別な統計ソフトなどを使わずに解くことも可能である。これが、欧州で1970年代から手法が使用されてきた所以である。またS字カーブが減少に転じる、いわば循環型・波動型の変化の分析や、言語変化の伝播の比率 p の上限値を予め設定する分析にもよく対応している。言語変化は化学変化とは異なり、短期間で

100%すべての人が新しい語形に移行するとは考えにくい事象や一旦進んだ変化が逆行する事象も多々考えられるので、②③のモデルのカーブを適用すれば結果の解釈も行いやすい。特に循環型・波動型のモデルはSPSSなどの統計ソフトのロジスティック回帰分析ではまだ対応していないものが多いので、循環型を考える場合には現状では③を用いることになる。

その一方、Altman学派のモデルには解法に限界もある。このモデルは基本的に比率 p と時間 t の2変数を前提としており、説明変数を2つ以上組み込む場合には自分で数学的な解法を考えなくてはならない。また式の特徴から比率 p が実際に0または1(=100%)になると数学的に解くことができない。言語変化で100%変化が完了するのを見届けるには非常に長い時間を要する場合もあるので、「限りなく100%に近づくが100%には達しない」データは言語学では現実的ではあるが、モデルの分析には不便な面もあろう。さらに予め比率 p の上限値を恣意的に設定することに違和感のあるデータも考えられる。

モデル④⑤のような形のロジスティック回

表1 S字カーブの各種モデルの比較

	①②③Altmanのカーブ	④⑤ロジスティック曲線 (下限値0、上限値1)	⑥ロジスティック曲線 (上下限値を推定)
背景となる主な分野	言語学	医学、心理学	人口学、経済学
主な解法	線形回帰に変形、 最小二乗法	線形回帰に変形、 最尤法	3群法による推定、 最小二乗法
解法ツール	なし(Excel等)	SPSS、R、エクセル統計等	R、エクセル統計等
独立変数の数	1つ	④は1つ、 ⑥は多変量に展開	1つ
上限値・下限値	予め設定する	下限値0、上限値1に固定	推定できる
循環型カーブ	対応	非対応	非対応
その他の注意点	$p = 0$ 、 $p = 1$ では不可		

帰分析はSPSSやRなどの統計ソフトに組み込まれており、自分で式を変形させて解く必要がない。解法には統計ソフトでは確率論に基づいた最尤推定法が使われることが多い。従って、その分析結果は単に過去の傾向の延長であるだけでなく確率的な実現の可能性を含んでいる。この点は最小二乗法を用いた重回帰分析と大きく異なる。

またロジスティック回帰分析では比率 p が0または1でも解くことができる。下限値と上限値は0と1に固定されているので、理論的にその範囲を逸脱しないという点で上限値・下限値を設定できない重回帰分析よりモデルに適用しやすい。しかし、例えば若年層にだけ広まりがちな流行語を広い世代で調査する事例のように、調査内容によっては伝播が100%に到達する前に飽和して減速する可能性が高いものに適用することは心理的な抵抗があるかもしれない。計算上はカーブの形状のどこかに適合して処理されるので、適用すること自体には問題はない。

さらに説明変数を多変量に展開できるので

比率 p と語形の伝播の時間 t のほかに、言語データの種別を示す質的変数をダミー変数として組み込んだり、調査地点の距離、調査年などほかの量的変数を組み込んだりすることもできる。この2つ目以降の量的変数からみても比率 p は単調増加の場合、S字を描いて増加していくことが確認されている（横山・真田2007b）。図6は仮想の方言データを2次元と3次元で表現した図で、2つ目の説明変数である調査地点間の距離でも「slow, quick, quick, slow」とよばれる伝播が成立することを示している。このように式を多変量に展開させることで今まで見えなかった角度から伝播の状況进行分析することは、言語変化の研究において非常に意味がある。

ロジスティック回帰分析のうちモデル⑥は、データの分布から下限値と上限値を推定できる。しかし推定する係数の数が多いため、直接、線形に変形して解くことができない。モデル④⑤と形が似ており統計ソフトを使って解くことができるが、線形に変形させて解く非線形回帰とは解法が異なり、修正指数曲線

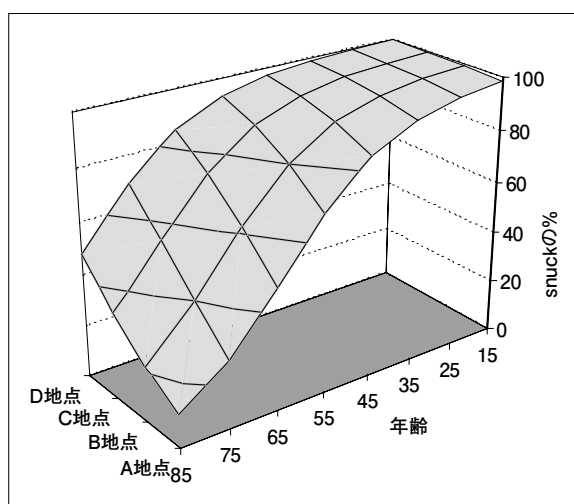


図6 仮想方言データの3次元のS字カーブ（横山・真田2007b）

(modified exponential curve) やゴンペルツ曲線 (Gompertz curve) と似た解法を用いる。

・理論の背景の思想的な問題

Altmann学派のモデルはもともと言語変化のダイナミズムを発想の基本にしているので、単なる手法の適用だけでなく言語変化の解釈の一部としてモデルを組み込むことができるという点で、言語学との親和性が高い。ロジスティック回帰分析は医学、心理学、生物学などの分野で分析ツールとして広く用いられているが、上述の通り言語学ではまだまだ一般的手法とはいえない。統計ソフトには組み込まれているとはいえ、計量言語学やコーパス言語学、自然言語処理などの分野を除くと統計ソフト自体は言語学では基本的なツールとはいえないからである。しかしそれだけでなく、発想の原点をどこに求めるかということも言語学では実は大きな影響がある。Altmann学派のモデルは言語変化を言語のダイナミズムによるものと考え、それを微分方程式で表現しようとしたが、これは思想の具現化の一手法であり言語学が古代の哲学に端を発したことを想起させる。ロジスティック回帰分析は現状では言語学にとってはツールの一つという位置づけになっている。新しい知見が得られる強力なツールではあるが、現状では言語学の思想に影響を与えるまでには至っていないようである。ロジスティック回帰分析は生物学や人口学の餌場理論などを基本にしているが、新語や新しい発音の定着が餌場理論を背景に持つと考えられるかどうかは今後の研究を待たなくてはならないだろう。

その一方で新しい可能性も感じられる。筆者はロジスティック回帰分析が確率論に基づ

く予測を行うことから、言語学でも確率論の考え方が導入できるのではないかと考えている。Altmann学派のモデルは最小二乗法を使った解法を行っているが、最小二乗法はいわば距離のモデルであり、現状の記述が重視されているといえる。確率論を視野に入れた分析により、言語学でも少ない実データから高い精度の予測が導ける可能性があり、記述的分析から仮説検証型の分析へと展開することが可能になろう。

5. おわりに

S字カーブを描いて時間とともに飽和に近づく言語変化のデータを、複数の変数で分析する多変量ロジスティック回帰分析の手法と適用事例を紹介した。また言語変化のS字カーブを分析する複数の手法について比較・検討を行った。ロジスティック曲線の研究は医学、心理学など言語学以外の分野ですでに展開されており、例えば項目反応理論にはS字カーブに上限値・下限値を設定したモデルがある^(註4)。またZにべき乗を組み込んだ高次式を代入すると循環型・波動型に対応したモデルになる。既存の調査データをこのような新しい手法で解析し直すことで新しい分析結果が得られる可能性もある。

一方で言語データを扱うことによる考慮も必要であろう。例えば複数の変数を設定する場合、その係数は調査年と被調査者の生年を同等の重みで扱えるのか、時代による社会的な影響をどのように数値化すればよいか、多変量S字カーブによって提示された結果は現実の言語生活の事象とどのように対応させて解釈すればよいか、定着と消失を繰り返す言語変化を表現するモデルはどのようなものが適しているのか、など他分野から手法を取

りこむだけでなく言語学になじむ手法にするための努力も必要であろう。

注

- (1) 変形の手順を参考のために以下に示す（横山・真田2007aより）。

$$p = 1 / [1 + A * \exp(-K * t)] \quad (\text{式1})$$

$A = \exp(\log A)$ を代入して整理すると

$$\begin{aligned} p &= 1 / [1 + \exp(\log A) * \exp(-K * t)] \\ &= 1 / [1 + \exp(\log A - K * t)] \\ &= 1 / [1 + \exp(-K * t + \log A)] \end{aligned}$$

ここで $a = -\log A$ とおくと

$$p = 1 / [1 + \exp(-K * t + a)] \quad (\text{式4})$$

となる。

- (2) アメリカでは、社会言語学者のLabov (1972) が1960年代から言語のフィールド調査にロジスティック回帰分析を適用している。この手法は、日本では日比谷潤子や松田謙次郎 (Hibiya, 1988; Matsuda, 1993) の研究が知られている。しかし、Labov学派は言語変化のS字カーブを近似・予測するためではなく、おもに、複数ある要因のうちどれが有意な効果を有するかを統計的に検定する目的でこの手法を用いており、Altmann学派とは関数の適用方法が異なっている。
- (3) この成果は、国立国語研究所のホームページ「敬語と敬語意識の半世紀—愛知県岡崎市における第三次調査—」(<http://www.kokken.go.jp/okazaki/>) に紹介されている。
- (4) 上述モデル⑥参照。このモデルは中村健太郎先生（埼玉学園大学）の御教示による。

参考文献

- Aitchison, J. (1991). *Language change: progress or decay?* 2nd ed. Cambridge: Cambridge University Press.
- Aitchison, J. (2000). *The seed of speech: Language origin and evolution*. Canto edition. Cambridge: Cambridge University Press.

- Altmann, G., von Buttlar, H., Rott, W., & Strauss, U. (1983). A Law of Change in Language. *Historical Linguistics. (Quantitative Linguistics. vol. 18)*. Bochum: Studienverlag Dr. N. Brockmeyer, pp. 104-115.
- Chambers, J. K. (1998). Social embedding of changes in progress. *Journal of English Linguistics*, vol. 26, pp. 5-36.
- Chambers, J. K. (2006). 国立国語研究所講演会資料
- Denison, D. (2003). Logist(ic) and simplistic S curve. *Motives for Language Change*. Cambridge: Cambridge University Press, pp.54-70.
- Hibiya, J. (1988). *A quantitative study of Tokyo Japanese*. Doctoral Dissertation, Dept of Linguistics, University of Pennsylvania.
- 橋本和佳 (2006). 「Logistic曲線による外来語増加過程のモデル化—大正から平成までの社説を用いて—」『計量国語学』、vol. 25-7, pp. 293-308.
- 井上史雄 (2000). 『東北方言の変遷』秋山書店
- 井上哲次郎、和田垣謙三、国府寺新作、有賀長雄 (1881). 『哲学字彙』東京大学三学部（1980年名著普及会刊の復刻による）
- 石川晃 (1993). 『市町村人口推計マニュアル』古今書院
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- 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) 秀英出版
- 国立国語研究所 (1957). 『敬語と敬語意識』(国立国語研究所報告11) 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) 秀英出版
- 国立国語研究所 (1983). 『敬語と敬語意識 岡崎における20年前との比較』(国立国語研究所報告77) 三省堂
- 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) 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]

pp. 63-74.

Matsuda, K. (1993). Dissecting analogical leveling quantitatively: The case of the innovative potential suffix in Tokyo Japanese. *Language Variation and Change*, 5, pp. 1-34.

南部智史 (2007). 「定量的分析に基づく「が／の」交替再考」『言語研究』, vol. 131, pp. 115-149.

真田治子 (2002). 『近代日本語における学術用語の成立と定着』 絢文社

Sanada, H. (2008). *Investigations in Japanese Historical Lexicology: Revised Edition*. Goettingen: Peust & Gutschmidt Verlag.

鐘水兼貴 (2006). 「東北・北海道における方言文法の共通語化過程」 *Linguistic Informatics VI, Linguistic Informatics and Spoken Language Corpora – Contributions of Linguistics, Applied Linguistics, Computer Sciences –*, Graduate School of Area and Culture Studies, Tokyo University of Foreign Studies, pp.365-381.

横山詔一、真田治子 (2007a). 「フィールド言語学にロジスティック回帰分析は寄与しうるか」『情報処理学会研究報告・人文科学とコンピュータ』 2007-CH-73, pp. 9-16.

横山詔一、真田治子 (2007b). 「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』 vol. 26-3, pp.79-93.

Yokoyama, Y., Sanada, H. (2007c). Multiple logistic regression analysis for formulating a change in language. *Proceedings of 5th Trier Symposium on Quantitative Linguistics*, p.22.

横山詔一、朝日祥之、真田治子 (2008). 「敬語意識の変化を予測する記憶モデル—多変量S字カーブによる解析—」『社会言語科学』 vol. 11-1, pp. 64-75.

横山詔一、真田治子 (2008). 「言語変化のS字カーブによる鶴岡市の共通語化予測」『日本語学会2008年度春季大会』 予稿集, pp.167-174.

Yokoyama, S., Wada, Y. (2006). A logistic regression model of variant preference in Japanese kanji: an integration of mere exposure effect and the generalized matching law. *Glottometrics*, vol. 12,

付記

この論文は、日本言語学会第137回大会公開シンポジウム「言語変化のモデル」(題目「言語変化のS字カーブの計量的研究—解析手法と分析事例の比較」2008年11月、於金沢大学)の発表の一部を元に加筆したものである。また本研究の一部は日本学術振興会科学研究費補助金(基盤研究(C)課題番号19520402研究課題名:『哲学字彙』にみられる近代学術用語の現代日本語への定着過程の検証)の助成を得ている。