

言語変化と社会環境

よこやま しょういち
横山 詔一 (国立国語研究所)

キーワード：普及のキャズム理論，言語変化のS字カーブ説，コーパスの異体字問題

1. はじめに

ここでは「社会環境」を次のように考える。

$$\text{社会環境} = \text{個人差} + \text{個人内差} + \text{時代差} + \text{地域差など} \quad (1)$$

個人差については、話者がいつ生まれたかの違い（世代差）などを取り上げる。個人内差は、同一話者内での加齢による差に着目する。時代差は、1950年以後の数十年間というスパンに限定して議論する。計量言語学の分野では Altmann ほか（1983）がペルシャ語に占めるアラビア語由来の外来語の割合について963年から1200年までの237年間にわたる変化を解析している。『Nature』にも言語変化研究の論文が掲載されることがあり、Lieberman ほか（2007）は800年前後から2000年前後までの約1200年間に生じた英語動詞活用形の変化について報告している。このような数百年単位あるいは千年単位の時代差を射程におさめた歴史言語学的研究の紹介は別の機会に譲る。地域差は、山形県鶴岡市における共通語化を検討する。

上記(1)式で個人差を世代差に、個人内差を加齢にした場合は、コウホート分析 (cohort analysis：コホート分析やコーホート分析などの表記もある) を適用する道が開ける。コウホート分析については松田（2017）が分かりやすい。また、言語変化にくわえて言語規範意識の変化に見られる時代差や地域差については「現代日本語史」という立場を打ち出した塩田（2011）が参考になる。

2. 普及のキャズム理論：個人差の分布法則と普及の壁

ルールを逸脱した表現のなかには一時的なブームとしていつの間にか消えていくものもある一方で、社会に許容されて定着していくものもある。両者の違いはどこにあるのだろうか？【シンポジウム概要から引用】

この問題を検討する際の手がかりとして、経営学のマーケティングやブランド構築に関する分野で Moore（1991）が提唱したキャズム (Chasm) 理論を紹介する。これは、革新的な商品が消えるか定着・普及するかについての説である。図1に示すように、この世の中には新しいモノが好きな人（イノベーターなど）と、そうでない人（レイトマジョリティなど）が一定の割合で存在すると考える。具体的には、イノベーターやレイトマジョリティなどといった個人差の分布は正規分布、つまりベル型カーブにしたがうと仮定されている。

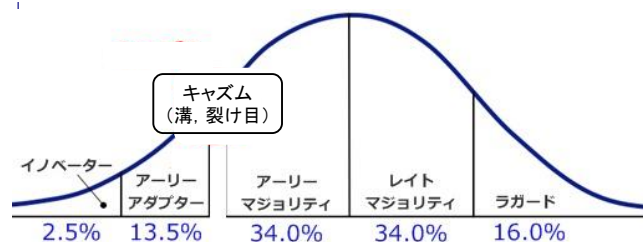


図1 キャズム理論

キャズムの意味は「溝(みぞ)」

とか「裂け目」である（本稿では「壁」と意識することもある）。図1の横軸は時間経過で、キャズム（普及率16%の壁）を越えると一気に普及が進む。逆に、その壁を越えられない商品は自然に消えていく。

キャズム理論は、普及学で Rogers (1962) が提唱した普及率の S 字カーブ説を基礎としている。この S 字カーブ説とよく似た考え方は、Aitchison (1981) による言語変化の S 字カーブ説に見られる。この説をもとに井上

(2008) が作成したのが図2で、規範意識の変化プロセスも示されている。図2の縦軸は逸脱表現の普及率で、横軸は時間経過（生年や世代）だと考えてもよい。この S 字カーブは、図1の普及率をイノベーターからラガード方向に累積したものとほぼ一致する。

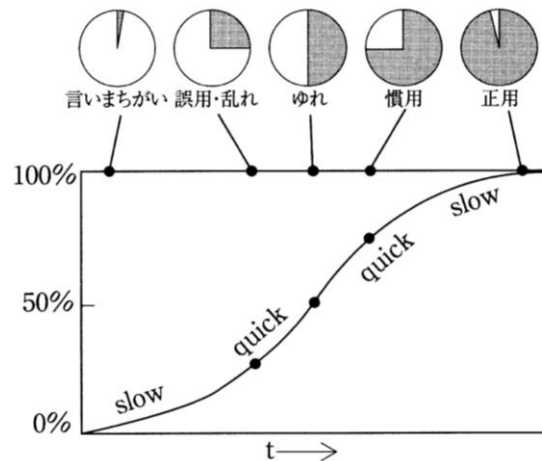


図2 言語変化の S 字カーブ（井上，2008 より）

3. 気づかない変化：個人内差は目立たない

さて、話し言葉や書き言葉のコーパスなどといった実際の言語データをもとにした近年の研究報告には、気づかない変化が身の回りに存在すること、そしてそれらを単なる誤用や逸脱表現として片付けられないことを示唆するものがある。【シンポジウム概要から引用】

新野 (2011) は、言葉に関心の深い人々にすら、変化していること自体ほとんど認識されず、言わば「深く静かに」変化が進行していると考えられるものがあると指摘し、そのような事例を「気づかない変化」と名付けた。そのヒントとなったのは、方言研究の分野で使われる「気づかない（「気づかれにくい」「気がつきにくい」などとも）方言」という概念であったという。

気づかない変化の研究に活用が期待できる誤用や逸脱表現に関する大規模調査は、文化庁の「国語に関する世論調査」などで毎年のように実施されているが今回は取り上げない。以下、誤用や逸脱表現ではないが、身の回りに存在する気づかない変化のヒントになりうる例として、山形県鶴岡市で実施された共通語化調査（以下、鶴岡共通語化調査）のうち成人の同一話者を41年間追跡した結果を報告する（横山ほか，2014）。

鶴岡共通語化調査は、国立国語研究所と統計数理研究所が共同で1950年から2011年まで61年間にわたり約20年間隔で4回実施したものである。データは国立国語研究所と統計数理研究所から順次公開されている（<http://www2.ninjal.ac.jp/longitudinal/>：最新版は ver.2.0）。そのデータを用いて1950年の第1回調査から1991年の第3回調査まで41年間にわたって3回連続で調査対象者になった話者を抽出した（53名：ちなみに2011年の第4回調査まで4回連続は4名）。さらに、鶴岡市で言語習得期あるいは言語形成期を過ぎた鶴岡ネイティブで、かつ第1回の時点で20歳以上だった38名（男性22名、女性16名）を選択し、今回の分析の対象とした。

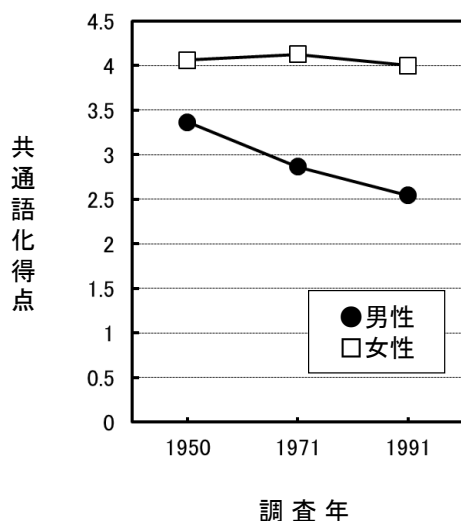


図3 音韻共通語化得点の平均値

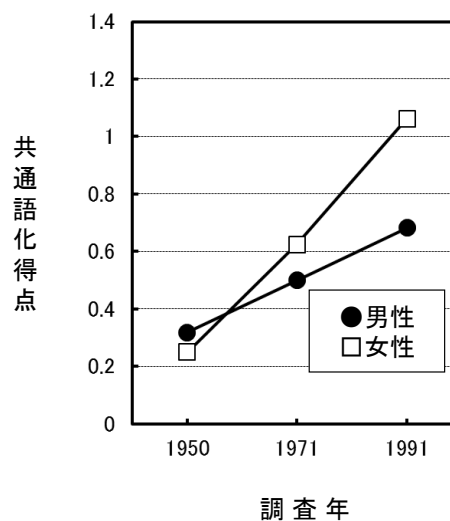


図4 アクセント共通語化得点の平均値

鶴岡共通語化調査では、音声に関する調査項目（全31項目）のうち「セナカ、ネコ、ハタ、カラス、ウチワ」の5項目については音韻とアクセントの両方のデータが記録されている。そこで、音声を聴き取って共通語と同じ項目がいくつあるかをカウントし、各話者の共通語化得点（0～5点）とした。図3に音韻共通語化得点の平均値を示す。分散分析の結果、性差がみられ（ $p<.01$ ）、女性が男性よりも共通語化得点が高かった。調査年の効果はなかった。図4はアクセント共通語化得点の平均値である。分散分析の結果、調査年の効果がみられ（ $p<.01$ ）、1991年の得点が1950年よりも高くなっていた。これは調査年が新しくなると共通語化が進むことを示している。性差はなかった。

上記の結果をまとめると、成人後も同一話者内で生涯にわたってアクセントの共通語化が進む一方で音韻は変化しないということになる。この現象に気づいている話者はきわめて少ないと推測する（ただし、その点の意識を話者に確認したデータはない）。アクセントは長期間にわたって少しずつ共通語化していくため、話者本人にくわえて周囲の人も気づかないという面があるだろう。また、すべての世代が同じ速度で共通語化していくと、あたかも自分は止まっている・変化していないように感じるということも見逃せない要因だと考えられる。

4. 言語変化の確率予測モデル：個人差＋[個人内差＋時代差]の効果

これまでに生じた言語変化のうちで定着してきたものは、われわれに変化の要因を読み解く手がかりを与えてくれる可能性がある。また、今起っている変化がこれからの日本語について予測するヒントになるかもしれない。では、言語変化を予測する手法にはどのようなものがあるのだろうか？【シンポジウム概要から引用】

ここでは、言語変化のS字カーブ説と関係が深いロジスティック回帰モデルを取り上げる。ガラス容器等のなかでショウジョウバエを飼育したり、シャーレのなかで細菌を培養したりして個体数を観察すると、序盤はゆっくり増加、中盤で急速に増加、終盤は

減速というように図2のようなS字カーブを描く。このS字カーブはロジスティック曲線あるいはシグモイド曲線と呼ばれ、個体密度が増えて生息環境が悪化すると個体数の増加にブレーキがかかる「密度効果」を表現したものと考えられている。

ロジスティック曲線は、正規分布を積分する普及率のS字カーブより数式が格段にシンプルである。両者の形状はほぼ同じで、差異は最大でも1%以内におさまることから、ロジスティック曲線は幅広い研究分野で頻繁に活用されている。

さて、先に紹介した鶴岡調査のアクセント共通語化データが多変量S字カーブを描いているか否かの判定は、ロジスティック回帰分析における統計的検定によるのが現在の定石である。ここでは「生年(世代)」は個人差に、「調査年」は個人内差(加齢)と時代差を合算したものと考える。この点でコウホート分析とは一線を画す。

図5に示す共通語化率は、先に述べたアクセント5項目から算出されたものである(横山・真田, 2010)。グラフ全体から第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、生年のほか調査年も説明変数に加えた多変量S字カーブが観測値と精度よく一致した。この結果は、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)を強く支持するものである。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、跳ね上がり現象を観測できないので、変化の速さを過小評価することになる。

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。個別に検討した結果においても、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。以上の実時間調査の結果から、アクセント5項目については「言語習得期を過ぎても生涯にわたって共通語化する」といえよう。この結果は先に示した図4の結果とも整合する。

図5を3次元的に表現すると図6のようになる。井上(1995)が提唱した「言語変化の水槽モデル」は図6ときわめてよく似た形状の曲面を描く。また、見かけ上の時間に

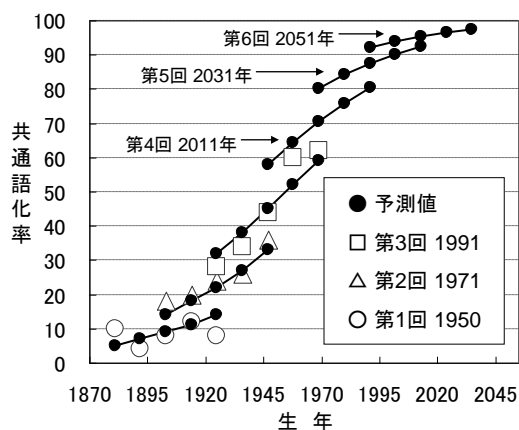


図5 アクセント共通語化の数値予測

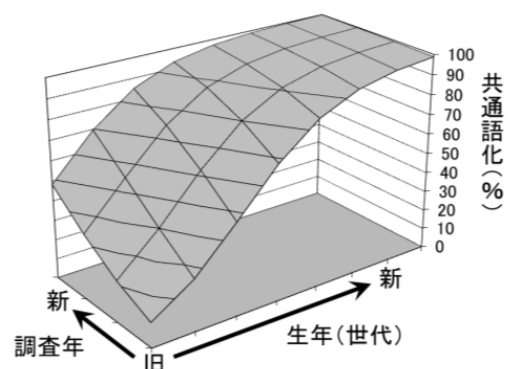


図6 多変量S字カーブの例

よる言語変化をロジスティック回帰分析で解析した最近の研究では久屋（2016）などがある。さらに、鶴岡共通語化のロジスティック回帰分析については、前川（2017）が最先端の統計的手法を用いた研究をおこなっている。

5. コーパスの異体字問題：文字政策の時代差を反映

ところで、言語変化や逸脱表現といった現象のありようを的確にとらえるためには、コーパスなどの言語データのとりあつかい方にも注意を払う必要がある。【シンポジウム概要から引用】

その典型例として、コーパスの異体字問題がよく知られている。たとえば、コーパスを用いて漢字頻度の経年変化を正確に捉えることはきわめて難しいことが示唆されている（横山ほか，1997）。その一例を示す。

異体字とは「桧—檜」のように読みと意味は同じで字体だけが異なる文字の集合を指す。異体字のバリエーション（変異）の豊富さは現代日本の文字生活になんらかの影響をもたらしているものと推測される。しかし、その詳細な実態はほとんど解明されていない（笹原ほか，2003）。そこで、現代社会の文字コミュニケーションに関する手がかりを得るために、異体字の使用頻度表を作成する場面を考えてみよう。たとえば「奈良県葛（ヒ）城市」の図 7A の字体と「東京都葛（L+人）飾区」の図 7B の字体が、新聞コーパスなどでそれぞれどのくらい出現するかを調べることは可能なのだろうか？

現在（2017 年）では、2007 年以前に作成したテキストファイルを IT 機器で表示・印字すると、図 7A の「葛（ヒ）」が図 7B の「葛（L+人）」に入れ替わって表示・印字されることが圧倒的に多い。本稿で「葛（ヒ）」や「葛（L+人）」という注記を付しているのは、この字体入れ替わり問題に対する予防策である。

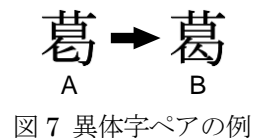


図 7 異体字ペアの例

日本語の異体字をめぐる IT 機器環境は、2007 年ごろから変化しつつある。たとえば、2006 年までは、一般的な文書作成ソフトで図 7B の「葛（L+人）」を使いたい場合であっても、原則として図 7A の「葛（ヒ）」しか使えなかった。Web での表示でも同様であった。ところが、2007 年 1 月に発売された Windows OS 「ビスタ（Vista）」で日本語フォントの約 150 字について字体が変更され、その影響で多くの IT 機器において図 7B の「葛（L+人）」が使えるようになった。一方、かつては使用できた図 7A の「葛（ヒ）」は IT 機器で使えない場合が増えている（横山，2016）。

マイクロソフト社がビスタで日本語フォント約 150 字の字体を変更した理由は、JIS 漢字規格の一つである「JIS X 0213:2004」に対応するためであった。周知の通り、JIS 漢字規格は漢字政策の一つとして強い影響力を持つ。この JIS X 0213:2004 とは、経済産業省が JIS 漢字（JIS X0213）の 160 字あまりについて印刷標準字体（いわゆる康熙字典体）にしたがって規格書の例示字形を 2004 年 2 月に変更したものを指す。IT 関連企業は、そのうちの約 150 字の異体字に対応したとされる。さらに、図 7B の「葛（L+人）」が改訂常用漢字表（2010）に追加されたことも社会に大きな影響を与えていると考えられる。このようにコーパスには文字政策の時代差がきわめて複雑な形で織り込まれている。コーパスを用いて漢字頻度の経年変化を調査する際はこの点に注意する必要がある。

6. 言語変化のダイナミクス

今後の研究には、当該の言語表現だけでなく、その言語表現が使用される環境や意味を成立させる諸条件などを視野に入れたダイナミクスについてもとらえることが求められるであろう。さらには許容する側である社会の「受け入れ態勢」にも注目すべきである。【シンポジウム概要から引用】

今回、なぜ言語変化が発動するのかという問題は紙幅の関係もあって議論しなかった。言語は合理的な方向に変化するといえるのであれば、合理的な形でスタートしないのはなぜかという疑問が残る。このような問題を含む言語変化のダイナミクスについては「見えざる手」理論にもとづく Fitch (2007) の解説などが参考になるかもしれない。

本研究は JSPS 科研費 17K18501 の助成を受けた。

おもな引用文献 (アルファベット順)

- Aitchison, J. (1991) *Language change: progress or decay?* 2nd ed. Cambridge: Cambridge University Press.
- Altmann, G., von Buttlar, H., Rott, W., & Strauss, U. (1983). A Law of Change in Language. In Brainerd, B. (Ed.), *Historical Linguistics. (Quantitative Linguistics. Vol.18)*. pp. 104-115. Bochum: Studienverlag Dr. N. Brockmeyer.
- Fitch, W. T. (2007) Linguistics: An Invisible Hand, *Nature* **449**, 665-667
- 井上史雄 (1995) 「音韻変化所要年数の規定要因—山添実時間調査—」『東京外国語大学論集』 **50**, 155-170.
- 井上史雄 (2000) 『東北方言の変遷』, 秋山書店
- 井上史雄 (2008) 『社会方言学論考—新方言の基盤—』 明治書院, 115
- 久屋愛実 (2016) 「見かけ上の時間を利用した外来語使用意識の通時変化予測」『日本語の研究』 **12**, 69-85
- Lieberman E., Michel JB., Jackson J., Tang T., & Nowak MA. (2007) Quantifying the evolutionary dynamics of language. *Nature* **449**: 713-716
- 前川喜久雄 (2017) 「鶴岡市共通語化調査データの確率論的再検討」『言語資源活用ワークショップ 2017 発表論文集』 163-180
- 松田謙次郎 (2017) 「岡崎敬語調査から学ぶ実時間調査の方法論的落とし穴」『敬語は変わる (大規模調査からわかる百年の動き)』 井上史雄 (編), 237-250, 大修館書店
- Moore, G. (1991) *Crossing the Chasm: Marketing in High-Tech Industries*. Harper Collins, New York
- 新野直哉 (2011) 『現代日本語における進行中の変化の研究—「誤用」「気づかない変化」を中心に—』 ひつじ研究叢書 (言語編) **93**, ひつじ書房
- Rogers, E. M. (1962) *Diffusion of Innovations*. New York, NY: Free Press.
- Sankoff, G. (2006) Age: Apparent time and real time. *Encyclopedia of Language and Linguistics*, 2nd edition, 110-115, Oxford: Elsevier.
- 笹原宏之・横山詔一・エリク＝ロング (2003) 『現代日本の異体字—漢字環境学序説—』 国立国語研究所プロジェクト選書 **2**, 東京: 三省堂.
- 塩田雄大 (2014) 『現代日本語史における放送用語の形成の研究』 三省堂
- 横山詔一 (2016) 「字体規範意識のデータサイエンス: 字体選好の地域差を探る」『漢字字体史研究 2: 字体と漢字情報』 石塚晴通 (監修) 高田智和・馬場 基・横山詔一 (編), 勉誠出版, 64-80
- 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**, 31-45
- 横山詔一・笹原宏之・野崎浩成・エリク＝ロング (1998) 『新聞電子メディアの漢字—朝日新聞 CD-ROM による漢字頻度表—』 国立国語研究所プロジェクト選書 **1**, 東京: 三省堂.
- 横山詔一・中村 隆・阿部貴人・前田忠彦・米田正人 (2014) 「成人の同一話者を 41 年間追跡した共通語化研究」『計量国語学』 **29**, 241-250