

# 言語の消失と獲得に関する環境・社会的要因

小野原 彩香

## Environmental and social factors associated with the gain or loss rate of language change

Ayaka ONOHARA

**Abstract:** Language change is affected by various factors such as environmental and social factors. In this study, linear mixed model was applied to estimate that the gain or loss rate of language change which associated with various factors in local area. This result indicates that rates of gain or loss words are lower in larger population area –that is, the large population cause hardly language change. This result showed opposing tendency to previous studies. So it's concluded that it is important to consider about the time and space scale.

**Keywords:** 言語進化 (language evolution), 言語地理学 (Linguistic geography), 人口統計学 (demography), 言語的多様性 (language diversity), 文化的多様性 (cultural diversity)

### 1. はじめに

近年, 言語変化の要因を特定するために, 多くの定量的研究が行われている. Bromham et al. (2015)では, 基礎語彙に注目し, 言語変化を語の消失と獲得に分けた上で, 地域の人口が多ければ, 新語の獲得率が高くなり, 少なければ, 語の消失率が高くなることを示している. Amano et al. (2014)では, 言語の消滅のメカニズムに注目し, その言語が話されている地域の面積の小ささと話者人口の少なさの両方が, 急速な話者数の減少と関連し, 話者の減少は経済成長に伴うものであることを示したほか, 高い地形的異質性 (topographic heterogeneity)の地域, 降水量の多い地域, 人口爆発地域が, 特に言語消滅のリスクにさらされていることを明らかにしている. Wieling

(2011)は, 人口規模があり, 平均所得が高く, 平均年齢が低いコミュニティは, 共通語に近づくことを明らかにしている.

このように, 言語変化の要因には, 様々なものがあると予測できる上に, 基礎語彙, 音韻, 文法, 語彙はそれぞれ異なる変化の過程を持つと予想されているため, 個々の変化がどのような要素と関係するかについて検証する必要がある. また, どのような時間スケールで, どのような空間スケールの問題を取り扱うのかについては, これまであまり議論がなされて来なかった. そこで本研究では, 先行研究で取り扱ってきた人口に関する要素 (人口, 人口増加率, 人口密度), 面積, 年齢構成 (高齢化率, 15歳未満率) と, 新たに土地の利用割合を加え, Bromham et al. (2015)に倣い, 言語の変化率を消失と獲得に分けた上で, 両者の関係を定量的に示し, 変化の原因とメカニズムに環境・社会的要因が及ぼす影響を明らかにすることを目的とする.

## 2. データの概要

### 2.1 対象地域と言語データ

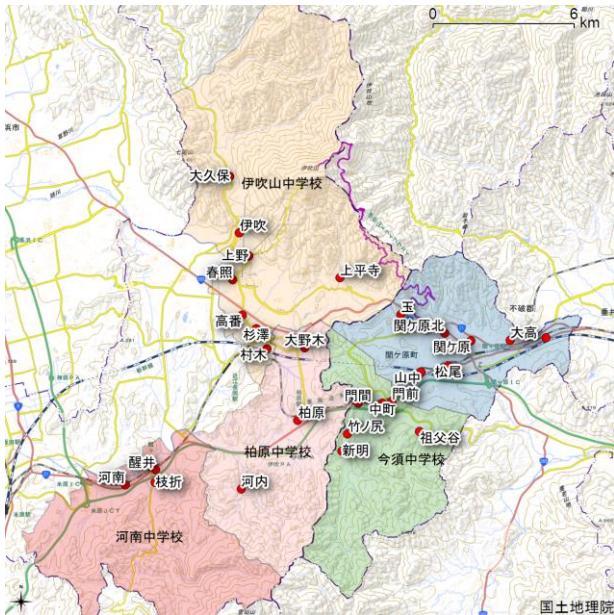


図 1 分析対象集落の位置関係<sup>1</sup>

対象地域は、滋賀県米原市の 3 学区 (柏原中学校, 伊吹山中学校, 河南中学校) と岐阜県関ケ原町の 2 学区 (関ケ原中学校, 今須中学校) である。言語データと環境・社会的指標との関係を検討するにあたって、対象地域中の 27 集落を使用した。該当集落は、図 1 にて示す。これは、対象地域のうち、祖父母と中学生世代の両世代において言語データが得られた集落のみを対象としているためである。

本研究で使用した言語データは、滋賀県米原市を中心に、文化・歴史の保全・教育・普及活動を行っている「はびろネット」が、2015 年 11 月 21 日～12 月 20 日に実施した方言アンケート調査の一部を使用している。はびろネットでは、米原市の 3 学区と関ケ原町の 2 学区の中学生男女と成人男女計 873 人 (対象地域全人口の 4.27%) からデータを収集したが、本研究では、中年層 (40 歳以上～60 歳未満) を除いた、中学生 544 人、60 歳以上 142 人のデータを使用している。調査項目数は、全 40 項目 106 語<sup>3</sup> であり、その内訳は、関西・近江系<sup>4</sup> の語 45 語、東京・若者言葉系の語 35 語、

美濃・尾張系の語 26 語である。

### 2.2 環境・社会的指標

言語に関する予想される環境・社会的指標については、土地利用における建物用地・森林・田の割合、人口(人; 2010 年)、人口密度 (人/km<sup>2</sup>)、人口増加率 (%), 高齢化率 (%; 2010 年), 15 歳未満率 (%; 2010), 面積 (m<sup>2</sup>) を使用した。建物用地・森林・田の割合については、国土交通省「国土数値情報、土地利用細分メッシュ」(平成 18 年度) (国土交通省, 2016) を使用した。人口、高齢化率、15 歳未満率、面積については、総務省 統計局 e-stat の平成 22 年国勢調査 (小地域)、年齢別 (5 歳階級、4 区分)、男女別人口 (総務省 統計局, 2016) を元にしている。

## 3. 方法

### 3.1 語の消失率・獲得率の算出

語の消失率・及び獲得率の算出は、以下の定義で行った。例えば、祖父母世代に伊吹集落では、美濃・尾張系の『アツコイ』(意: 分厚い)を 25% の人が伝聞・使用<sup>5</sup>しているが、中学生世代では 12.5% の人が伝聞・使用している。このとき、率が減少しているため、これを消失率とみなし、 $0.25 - 0.125 = 0.125$  ポイントの消失率と定義し、率が増加している場合を獲得率と定義する<sup>6</sup>。以上の定義にて 27 集落 106 項目すべてに対して、祖父母世代から中学生世代へかけての語の得失率を算出し、項目を関西・近江系、東京・若者系、美濃・尾張系の 3 タイプに分けた上で、それぞれの平均値を算出した。この方法で最終的に、27 の各集落で、各タイプの項目ごとに消失率・獲得率を得た。

### 3.2 言語と環境・社会的指標の関係についてのモデル

上記の過程で得られた語の消失率と獲得率が、環境・社会的要因とどのような関係にあるかについては、中学校の学区をランダム効果とする線形混合モデル (Linear Mixed Model: LMM) を用いて評価した。なお、計算に当たっては、R のパッケ

表1 各項目のタイプ別消失・獲得率と環境・社会的要因についての線形混合モデル(AIC最小)

変化率	説明変数の係数										Intercept	AIC	BIC	logLik	deviance	df.resid	
	経度	緯度	建物用地	森林	田	人口	人口密度	人口 増加率	高齢化率	15歳未 満率							
関西消失率	-	-	-	-	-	-0.624	-	0.299	-	-	-	<0.001	70.05	76.53	-30.03	60.05	22
東京消失率	0.410	-	-1.107	-	-	-0.271	1.020	-	-0.829	-0.701	-	-0.014	63.30	74.96	-22.65	45.30	18
美濃消失率	0.387	-	-	-	-	-0.423	-	0.240	-0.244	-	-	<0.001	75.27	84.35	-30.64	61.27	20
関西獲得率	-	-	-	-	-	-0.495	-	-	-	-	-	<0.001	76.03	81.22	-34.02	68.03	23
東京獲得率	0.346	-	-	-	-	-0.748	-	-	-0.225	-	0.275	<0.001	66.12	75.19	-26.06	52.12	20
美濃獲得率	0.446	-0.27	0.514	-	-	-0.859	-	-	-	-	0.773	<0.001	69.82	80.19	-26.91	53.82	19

ージ lme4 の関数 lmer (Douglas, 2016)を使用し、適切な説明変数のみをモデルに組み込むために、赤池情報量基準 (Akaike's Information Criterion: AIC) (Akaike, 1974) の値にてモデル選択を行った<sup>7</sup>。

また、今回使用するデータに関しては、説明変数の Moran' I および Geary's C、線形回帰モデルでの誤差項の Moran' I、説明変数と誤差項のラグランジュ乗数検定<sup>9</sup>において、空間的自己相関は見られなかった。さらに、使用データは、個体(個人)差についていわゆる反復<sup>10</sup>、場所(集落)差については疑似反復であり、各測定値の分布から、正規分布に従うと想定できたため、線形混合モデルを使用し、説明変数の単位を揃え評価をしやすくするために、あらかじめ標準化したデータをモデルで評価した。

#### 4. 結果

表1<sup>11</sup>は、関西・近江系項目、東京系項目、美濃・尾張系項目それぞれについて、獲得率と消失率ごとに、環境・社会的指標との関係を示す AIC の最も低いモデルを一覧<sup>12</sup>にしたものである。表の説明変数の係数を見ていくと、どの系統の項目においても、獲得率、消失率共に人口が関わっていると予想されている。そして、係数が負の値を取ることから、人口が多くなれば多くなるほど、語の獲得率、消失率共に低くなることが分かる。これは、地域の人口が多くなればなるほど、変化が起きにくくなることを示している。また、消失率、獲得率に注目すると、消失率に特有の要素は、人口増加率<sup>13</sup>であり、獲得率に特有の要素は面積である。人口が増加しているような場所では、言

葉が失われやすく、集落面積が大きければ、新しい言葉を得やすいと解釈できる。さらに、3つのタイプに注目すると、東京・若者系項目と美濃・尾張系項目で、経度が関係しており、経度が大きくなる、すなわち関ヶ原側では、語の消失率・獲得率が共に高いことが分かる。

#### 5. 考察

以上、言語項目を3タイプに分けて、言語の消失率と獲得率に関わる環境・社会的要素についてそれぞれの影響を評価した。いずれのタイプにおいても、人口が多く、高齢化率が高いと、使用項目の変化が起こり難いことが明らかとなった。これは、Bromham et al. (2015)の、地域の人口が多ければ新語の獲得率が高くなり、少なければ語の消失率が高くなるという結果のうち、人口が多い場合の結果を支持せず、逆のパターン、人口が多いと、獲得率は高くならないことを示している。これは、空間スケールもしくは時間スケール、あるいは、調査項目の問題に帰結するのではないかと考えられる。Bromham et al. (2015)では、人口は270人~199000人の範囲にあり、本研究は、56人~3521人の範囲にある。また、Bromham et al. (2015)では300年~3000年の基礎語彙の消失・獲得を問題にしているのに対し、本研究では、約50年の時間経過の中での語彙・文法面の消失・獲得を問題にしている。このため、今後、空間・時間・対象の3スケールで、言語変化と関わる項目の影響がどのように変化するか、何らかの閾値があるかを連続的なスケールにて探る必要がある。

また、関ヶ原では、東京系と美濃系において言葉の交替が早いと予想される。この原因について

は、さらなる検討が必要である。

さらに、ある程度人口が多い場所では変化が起き難くなるが、人口密度や人口増加率が高ければ使用項目の消失は起こりやすくなる。これは、Amano et al. (2014)の人口爆発地域において言語消滅が起こりやすいという結果の人口爆発地域を人口増加率の高い地域と読み替えれば一致する。

共通語化に着目すると、人口が少なく、高齢化率が低くなると、東京系の獲得率が高くなる。これは、Wieling (2011)は、人口規模があり、平均所得が高く、平均年齢が低いコミュニティは、共通語に近づくという結果のうち、平均年齢の低さのみを支持する結果であり、人口規模については、ここでも逆の結果を得ている。

## 6. 結論

本研究では、滋賀・岐阜県境域の時系列な言語データを用いて、言語の消失と獲得にどのような環境・社会的指標が関係するのかを見てきた。本研究では、人口が多く、高齢化率が高いと、使用項目の変化が消失・獲得共に起こり難いことを示した。また、人口密度や人口増加率が高ければ消失率が高くなり、面積が大きければ獲得率が高くなる。Bromham et al. (2015)では、人口が大きい場合に、語の獲得率が高くなるという結果であるが、本研究では逆の結果が得られた。これは、時間・空間・対象項目の3つの観点からさらに詳細な検討を加えていく必要がある。

## 謝辞

本研究における言語データは、はびろネット（滋賀県米原市柏原）が、米原市「まいばら協働事業提案制度」にて行った方言調査アンケート（調査総数809人）の一部を使用したものです。データ分析並びに学術論文作成について、ご快諾くださったはびろネット事務局の方々に感謝申し上げます。また、杉崎好洋氏には、美濃・近江方言について、日常の議論を通じて多くの知識や示唆をいただき、文献資料を譲与・貸与頂いたばかりでなく、米原、関ヶ原地域の巡査にお誘いいただき、当該地域の詳細を知るきっかけを与えてくださいました。ここに感謝の意を表します。

そして、調査対象地区の各自治会および調査対象者の皆様、データ入力をなさった方々に心から感謝します。

本研究は、JSPS 特別研究員奨励費 13J03462 の助成を受けたものです。

<sup>1</sup> 図1は国土地理院 地理院タイル（標準地図）および国土数値情報 中学校区データ（データ作成年度:平成25年度）を加工して使用。

<sup>3</sup> 例 美濃・尾張系 名詞 チヨンチヨン (鉛筆が尖っている様)。

関西・近江系の語45語: 文末詞2, 名詞15, 文法7, 待遇6, 動詞5, 副詞2, 助動詞2, 形容詞5, 句1

東京・若者言葉系の語35語: 文末詞1, 名詞10, 文法5, 待遇0, 動詞13, 副詞2, 助動詞0, 形容詞4, 句0

美濃・尾張系の語26語: 文末詞2, 名詞10, 文法1, 待遇4, 動詞5, 副詞1, 助動詞2, 形容詞1, 句0)

<sup>4</sup> 例えば、「雨だ」は、東京系であると同時に、尾張系でもある。美濃で「雨や」→「雨だ」と世代交代を経て変化しつつあるのは、名古屋の影響だと見える。このように言葉の由来や受け入れ方には複雑な背景がある。ここで、○○系としたのは便宜的なものである。

使用については、全体の何割の人が使用しているかの率。伝聞経験（「聞いたことがあるが使用しない」）は使用の半分の率として計算。

<sup>6</sup> 獲得率: 例えば、柏原の『あかんヤン』の伝聞・使用率は、祖父母世代0.75、中学生世代0.875なので、獲得率0.125。

<sup>7</sup> Rのパッケージ MuMin(Bartoň, 2016)のdredgeを使用。

<sup>9</sup> Rのパッケージ spdep (Bivand R. 2016)を使用。

<sup>10</sup> 同じ人に複数回アンケートを取っていないので反復(replication)、同じ地域から複数人にアンケートを取っているので疑似反復(pseudo replication)。

logLik: 最大対数尤度(logL), deviance: 逸脱度(D=-2logL), df.resid: the residual degrees-of-freedom extracted from a fitted model object

<sup>12</sup> 各タイプ上位5つのモデルにおいても、ほぼ同じ変数が選択される。

<sup>13</sup> 説明変数同士の相関係数。建物用地-人口密度: 0.887, 建物用地-人口: 0.564, 人口密度-人口: 0.489, 人口密度-人口増加率: 0.444, 人口増加率-建物用地: 0.265, 面積-人口: 0.185, 人口増加率-人口: 0.111, 面積-人口増加率: -0.292, 面積-建物用地: -0.298, 面積-人口密度: -0.363

## 参考文献

- Bromham, L., Hua, X., Fitzpatrick, T. G., & Greenhill, S. J. (2015). Rate of language evolution is affected by population size. *Proceedings of the National Academy of Sciences*, 112(7), 201419704.
- Amano, T., Sandel, B., Eager, H., Bulteau, E., Svenning, J.-C., Dalsgaard, B., … Sutherland, W. J. (2014). Global distribution and drivers of language extinction risk. *Proceedings. Biological Sciences / The Royal Society*, 281(1793), 20141574-.
- Wieling, M., Nerbonne, J., & Baayen, R. H. (2011). Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PLoS ONE*, 6(9).
- 国土交通省, 国土数値情報 土地利用細分メッシュ <http://nlftp.mlit.go.jp/ksj/>, 閲覧日 2016/07/01.
- 政府統計の総合窓口 e-Stat, 総務省統計局 <https://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do>, 閲覧日 2016/07/01.
- Douglas B. 2016 lme4: Linear Mixed-Effects Models using 'Eigen' and 'S4': R package. <http://cran.r-project.org/web/packages/lme4/index.html>. 閲覧日 2016/07/01.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.
- Bartoň K. 2016 MuMin: multi-model inference: R package. <http://cran.r-project.org/web/packages/MuMin/index.html>. 閲覧日 2016/07/01.
- Bivand R. 2016 spdep: spatial dependence: weighting schemes, statistics and models: R package. <http://cran.r-project.org/web/packages/spdep/index.html>. 閲覧日 2016/07/01.