

静岡方言「まめったい」に見られる意味変化：
多項分布型レジームスイッチング検出を用いて

Semantic Change in Shizuoka Dialect *Mamettai* : Analyzed by Multinomial Distribution Type Regime Switching Detection

峯尾海成[†] 佐藤道大[†] 山岸祐己[†] 谷口ジョイ[†]
Kaisei Mineo Michihiro Sato Yuki Yamagishi Joy Taniguchi

1. はじめに

本研究は、静岡方言「まめったい」に関する大規模調査を行い、その結果を多項分布型レジームスイッチング検出および多群出現順位統計量に基づく時系列データの変換によって分析することで、過去に生じた意味変化とその生成時期についての特定を試みたものである。

言語・方言の衰退、消滅は多くの国や地域でみられており、日本においても消滅の危機にある方言（以下、危機方言）が複数存在している。国内の危機方言は、主に琉球語諸方言について、国立国語研究所を中心に記録・保存が進められているが、静岡方言においては、メディアの流入や交通環境の改善によって共通語化が進み、話者の数が激減しているにもかかわらず、データの蓄積は進んでいない。

本研究では、静岡方言「まめったい」の使用・理解について大規模調査を行い、それらを年代別の時系列データとして分析することで、過去に変化が生じた時期を特定することを試みる。なお、本調査で扱う時系列データは、変化が複雑であり、情報量も多く、そのまま可視化しても解釈が困難である。よって本研究では、データを分かりやすい形で可視化する「多項分布型レジームスイッチング検出」を用いて、データの単純化を行う。また、多群出現順位統計量により時系列データの変換を行い、長期的な意味変化の傾向を捉える。

2. 関連研究

2.1 静岡方言「まめったい」について

静岡で用いられる「まめったい」という方言には複数の意味が存在する。全国方言辞典[1]によれば、(1) よく働く、(2) 達者だ、という意味があり、それぞれ「隣の嫁はまめったいね（隣の嫁はよく働くね）」、「あのおじーさんまめったいね（あのおじーさんは達者だね）」という用例が付記されている。本調査の実施にあたり、静岡方言を母方言とする話者を対象に予備調査を行ったところ、上記以外にも、(3) (子どもが) 落ち着きがない、(4) 几帳面だ、細かい点まで徹底する、といった意味で使用されていることが示唆された。また、「まめったい」をどのような意味で使用するかには、個人差があり、年代や地域が影響を与えている可能性について考慮する必要があると考えた。以上より、本研究では以下 2 点の仮説を検証する。

- (1) 静岡方言「まめったい」には複数の意味が存在するが、その使用には年代差がある。

- (2) 静岡方言「まめったい」は過去に意味変化が生じている、あるいは変化が進行中である。

2.2 多項分布型レジームスイッチング検出

多項分布型レジームスイッチング検出は、タイムラインを生成することで、複雑に変化するデータを単純化する手法である。本研究では、データが多項分布に従うと仮定し、多項分布型レジームスイッチング検出をデータの単純化手法として用いることで、静岡方言「まめったい」において、過去に生じた意味変化とその時期を推定する。既存の類似手法[2]では、単一情報の傾向変化を可視化することに特化しているが、山岸ら[3]によって提案された手法においては、複数情報を扱うことを前提としており、変化が生じた期間の把握が容易であることが示されている。

2.3 多群出現順位統計量に基づく時系列データの変換

データカテゴリの時系列的変化を明示し、それらを複数カテゴリ間で比較するため、出現順位を用いた統計量によるデータ変換[4]を行う。この手法は、Mann-Whitney の U 検定を基盤とし、多群を扱えるよう拡張したものであり、データの出現頻度の傾向変化を z-score として表現する。静的な分析手法による指標を、動的な視点で捉えられるよう可視化するため、長期的な変化が捉えやすいと言える。また、各カテゴリの z-score は、他のカテゴリすべてを基準としているため、複数間のカテゴリにおける比較が容易である。

3. 研究手法

3.1 まめったいの大規模調査

3.1.1 調査対象者

本調査は、出生から中学校卒業時まで継続して静岡に居住していた、静岡方言を母方言とする 1,544 名（男性 599 名、女性 945 名）を対象としている。県東部、及び県西部は若年層の回答者が多く、県中部に関しては、1950 年代～1970 年代生まれの回答者が多い。また、井川は、高年層の回答者が多いが、これは井川の高齢化率が 62.37% という高い値を示していることに関係している。調査協力者の内訳は表 1 に示す。

[†] 静岡理科大学
Shizuoka Institute of Science and Technology

表 1 調査対象者の内訳

	県西部	県中部	井川	県東部	総計
10 歳代	22	30	4	27	83
20 歳代	111	182	2	117	412
30 歳代	55	140	3	71	269
40 歳代	55	208	1	71	335
50 歳代	40	184	3	44	271
60 歳代	15	93	5	6	119
70 歳以上	4	25	23	3	55
計	302	862	41	339	1544

3.1.2 調査対象者

本調査は、質問紙、あるいはウェブ調査により、性別、生年、出身地、家庭内における方言使用の有無、「まめったい」の使用・理解について回答を依頼した。出身地については、中條[5]に基づき、地図を示した上で、東部・中部・西部・井川の 4 地域を選択肢として設けた。また、「まめったい」を使用、あるいは理解すると回答した協力者には、その意味を選択してもらい（複数回答可）、選択肢以外の意味で使用する場合は「その他」の欄への記述を依頼した。さらに、(1) よく働く（体を動かす）、(2) 達者だ、(3) 落ち着きがない、(4) 勤勉だ、細かい点まで徹底する、という 4 つの用法について、その用例と意味を示した上で、「使用する」「使用しないが理解できる」「聞いたことはあるが、これまで理解できなかった」「聞いたことがない」という 4 つの選択肢を設けた。最後に、「まめったい」を用いた短文作成を依頼し、上記 (1) から (4) に基づく意味分類を行った上で、肯定的な意味で使用されているか、あるいは否定的に用いられているかを分類した。

4. 分析結果

図 1, 2, 3 は、「まめったい」の使用・理解に関するデータを、それぞれ異なる手法で可視化したものである。図

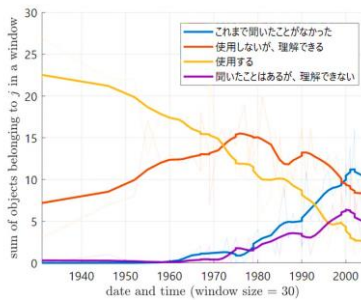


図 1 一般的な可視化 (30 観測ごとのカテゴリの数)

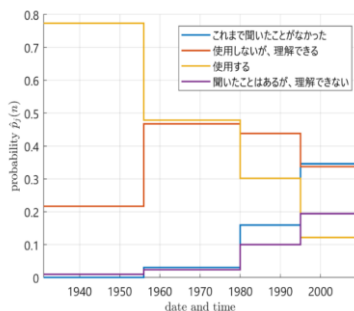


図 2 多項分布型レジームスイッチング検出

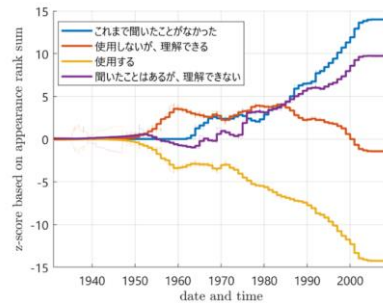


図 3 多群出現順位統計量

1 に、一般的な可視化を示す。質問紙調査によって得られた回答は独立したデータであるため、30 観測ごとに数え上げている。若年層ほど「使用する」、「使用しないが、理解できる」との回答が減少傾向にあるが、変化の時期を推定するのは困難である。図 2 は、多項分布型レジームスイッチングによるタイムラインを示したものであり、1956 年、1980 年、1995 年（生年）でデータ構造に変化が生じていることが示唆されている。変化点数の推定については、最小記述長というモデル選択基準を用いている。これは、膨大で複雑なデータから規則性を抽出する際に用いられる技法であり、これ以上生年の区切りを増やすと情報が過剰となり、また逆に区切りを少なくすると、データのもつ情報の表現方法としては過度な単純化が生じるため、上記が最も効率の良い分け方と言える。図 3 では、多群出現順位統計量の z-score により、「まめったい」の使用・理解の度合いについて、長期的な変化傾向とカテゴリ間の勢力関係を定量的に示した。

5. 考察及びまとめ

本研究では、静岡方言「まめったい」の意味変化について、多項分布型レジームスイッチング、及び多群出現順位統計量によって、周期的時系列データを用いたデータの単純化及び可視化を試みた。本手法により、語の意味変化が生じた時期を推測した結果、「まめったい」の使用・理解においては、1956 年、1980 年、1995 年（生年）でデータ構造に顕著な変化が見られた。また、特に 1980 年以降生まれでは、z-score の継続的な現象が観測され、p 値が 0 に近似していることから、この減少傾向は偶然と言えない。よって、「まめったい」の使用・理解度の低下は、今後も継続することが推測される。

本研究で扱ったデータは、一般的な可視化からの分析が困難であったが、本研究で用いた手法により、語の意味変化の時期の推定が可能であった。これらの手法のさらなる活用が期待される。

参考文献

- [1] 東条操“全国方言辞典”，東京堂，(1951).
- [2] J. Kleinberg. Bursty and hierarchical structure in streams. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD-2002)*, pp. 91–101, 2002.
- [3] 山岸祐己, 岩崎清斗, 齊藤和巳. “多項分布型レジームスイッチング検出による周期的時系列データの単純化”, 情報処理学会論文誌数理モデル化と応用(TOM), Vol. 11, No. 2, pp. 99–106, (2018).
- [4] 山岸祐己, 岩崎清斗, 齊藤和巳. “多群出現順位統計量に基づく時系列データの変換”, 情報処理学会論文誌数理モデル化と応用(TOM), Vol. 11, No. 1, pp. 45–52, (2018).
- [5] 中條修“静岡方言の研究”吉見書店, (1982).