

『日本言語地図 第3集』データを使った言語的地域性の解析

熊谷康雄 (国立国語研究所情報資料研究部)

1. はじめに: 近年, 旧来の方言区画の方法に対して, データ処理による客観的な手法として計量的方言区画論が説かれ, 様々な試みがなされている (例えば, 井上 (1986, 1996), 小林 (1988) を参照). 柴田 (1964) は方言区画の概念を整理し, 方言区画とは方言による地域の分割であるとし, 方言の分布と区別して定義した (なお, 熊谷 (1996) では, この概念規定について再考した). 本発表はこの観点から出発して開発を続けている計量的方言区画のための「ネットワーク法」を『日本言語地図 第3集』データへ適用する試みである.

2. データ: 『日本言語地図』(全6巻)は国立国語研究所が1957~1964年に, 全国2400地点を対象に行った言語地理学的な調査の成果である. 調査員による臨地調査であり, インフォーマントは各調査地点につき1名, 1903年以前出生の男子, インフォーマント自身の使用語を対象とし, などなぞ方式の質問により調査し, 総調査項目数は285であった. 回答には複数回答がある. 全6巻のうち, 第3集が機械可読化された (国立国語研究所1986). 第3集は語彙項目のみであり, 2400地点, 45項目のデータがある. このうち, 沢木 (1988) が孤例の分析のために選択した27項目を対象とする. これは項目により調査地点数にばらつきがあるので, 条件を揃えるために, 調査地点数が総地点数2400地点に近い項目を選んだものである. ネットワーク法では個々の回答語形を項目として扱う. この意味では, 調査地点数2400, 項目数 (異なり語形数) 4151 (調査項目27) が処理対象となったデータである.

なお, 27項目とは, 頭, 旋毛, 禿げ頭, 顔, 頬, 目, 眉毛, 麦粒腫, 唇, 舌, 唾, 涎, 親指, 人指し指, 中指, 薬指, 小指, しもやけ, 腫, ほくろ, 女, 風, 竹馬, お手玉, 鬼ごっこ, かくれんぼ, 肩車である.

3. ネットワーク法の適用: 「ネットワーク法」は言語地理学的な調査結果のデータから地点同士の類似の度合を計算し, 地図上にネットワークの形で表示することによって, 地理的平面上に言語的に類似した地点が広がっている様子を観察できるようにするものである.

3.1 共有度: 任意の2地点間の類似の度合いを測る尺度のひとつとして, 共有度を定義する. これは2地点が共有する項目 (ここでは回答語形) の数である. ネットワーク法は全ての地点の組み合わせについて共有度を計算し, これによって得られるマトリックスを処理する. (なお, 地点同士の類似の度合いの尺度として, 他に, 各地点が示す共有度のパターン相互間のユークリッド距離をd法と称して試みている. 今回のデータに関してもd法による処理を進めている.)

3.2 NT-1(r)n: 共有度から見て, 一定の値(Lcond)以上似ている地点のペアを地図上で線を引き結ぶことによって得られるネットワーク図を基本とし, この値Lcondを条件を変化させながら, ネットワーク図を観察することにより, 言語的に類似した地点からなる地点のグループがどのように広がっているのかを観察するものである. 観察には類似性と地理的連続性に関して読み取りのためのルールを用意している. 図1に一覧表示のネットワーク図を, 図2, 図3にはLcond=22と21の場合のネットワーク図を載せる. なお, ここでは北海道との関係を示す枝はネットワークの表示の際に表示を抑制してある. 例えば, 関東や関西を中心とした地域の広がり, 東西の方言の境界にある中部地方が中部地方としてもまとまりを見せつつ, 全体としては関東地方の広がりの上に乘ってくる様子や, 北陸地方の東西の境界としての様子など, よく観察される.

3.3 Delaunay net 上での観察: 地図上で, 地点間の類似関係が地面に沿って推移する様子を観察するために, 調査地点の隣接地点を求め, 隣接地点のネットワーク上での類似関係の観察を考えた. 実現のために, 隣接する地点というものの求めかたのひとつとして, 調査地点に対して, Delaunay net の計算をした (図4,6). これによって得られた「隣接地点」同士を結ぶ枝に共有度の値を付与し, ネットワーク図として共有度の値の分布を視覚的に表示した. 図7,8,9はLcondを満たす枝のみを表示した. 図5には共有度を5段階にカテゴライズして一枚にまとめたものを示す. 地域的なつながり方が観察でき, 3.2節の表示と直接突き合わせができる.

4. おわりに

日本言語地図第3集のデータにネットワーク法を適用し, 言葉から見た地域性の解析に有効な観察ができた. さらに, 他の地理的情報や解析結果との突き合わせをしながら, 手法の開発を進める予定である. 最後に, 日本言語地図第3集の機械可読データを作成された佐藤亮一, 沢木幹栄, 小林隆, 白沢宏枝の諸氏に感謝します.

文献

- 井上史雄 (1986) 文法現象による計量的方言区画, 言語研究89号
- 井上史雄 (1996) 方言分布の多変量解析法と分布重心法. 『言語学林 1995-1996』三省堂
- 熊谷康雄 (1993) 数量化理論 III 類による方言区画とネットワーク法 —「井上史雄『文法現象による計量的方言区画』のネットワーク法による再分析—」日本方言研究会第57回研究発表会発表原稿集
- 熊谷康雄 (1996) 方言区画の方法における地理的な次元. 『言語学林 1995-1996』三省堂
- 国立国語研究所 (1968) 『日本言語地図 第3集』大蔵省印刷局 (縮刷版1983)
- 国立国語研究所 (1986) 『方言研究と電子計算機』昭和60年度科学研究費補助金 (一般研究B) 研究成果報告書
- 小林隆 (1988) 方言の分類. 『日本語百科大事典』, 大修館書店
- 沢木幹栄 (1988) 『日本言語地図』の語形の数量的性質. 国立国語研究所『方言研究法の探索』, 秀英出版
- 柴田武 (1964) 方言区画とは何か. 東条操監修, 日本方言研究会編『日本の方言区画』, 東京堂 (柴田武『方言の世界 ことばの生まれるところ』, 1978, 平凡社に再録)

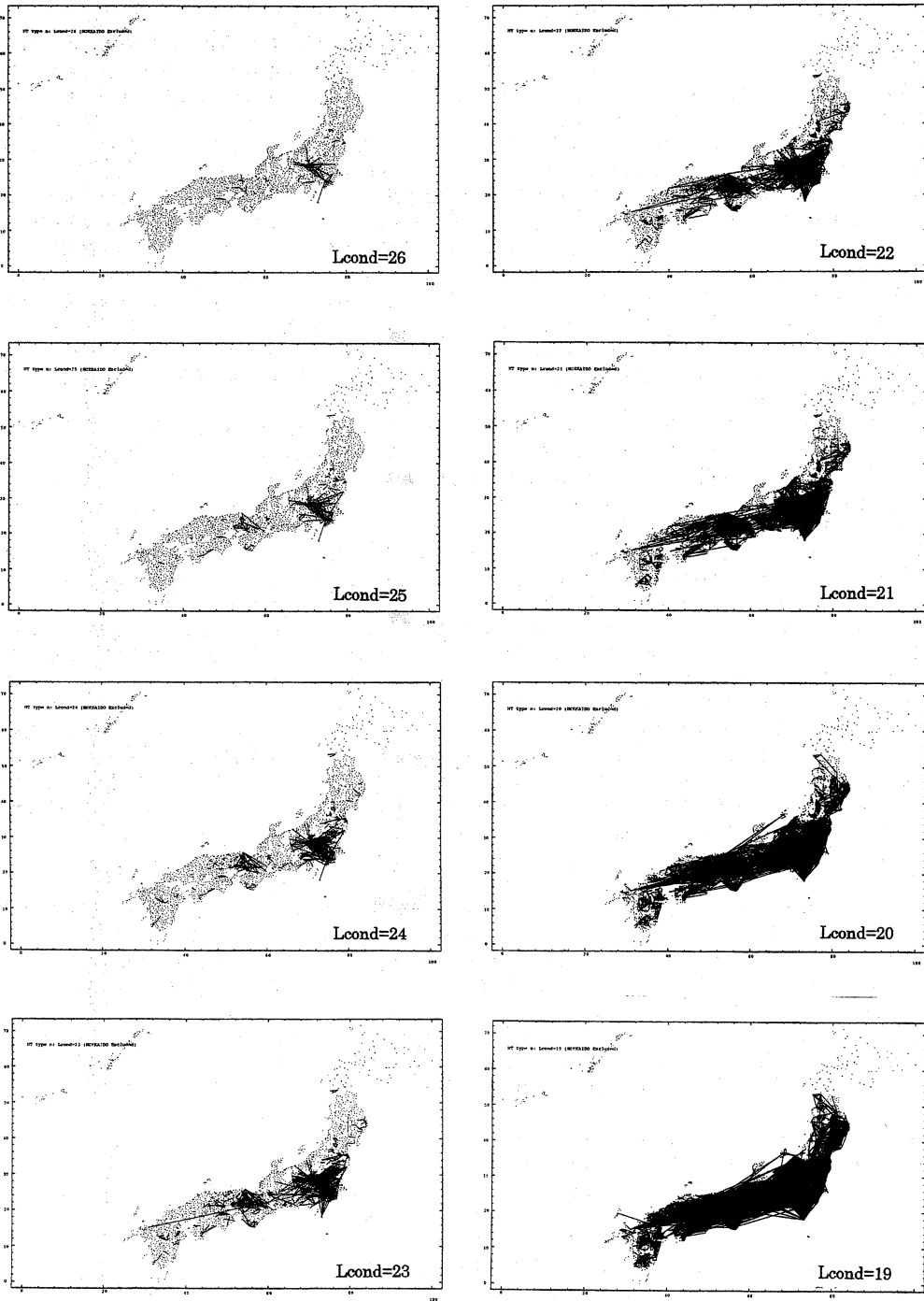


図1 ネットワーク図 (n法 Lcond=26~19) 一覧表示 (北海道との関係の表示を抑制)

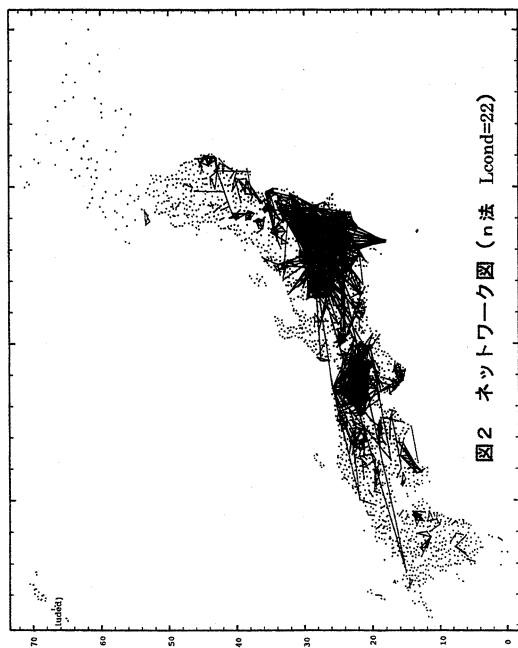


図2 ネットワーク図 (n法 Lcond=22)

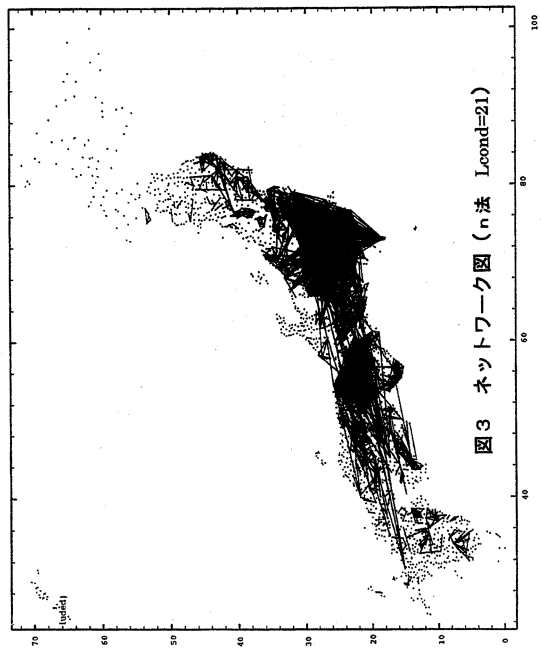


図3 ネットワーク図 (n法 Lcond=21)

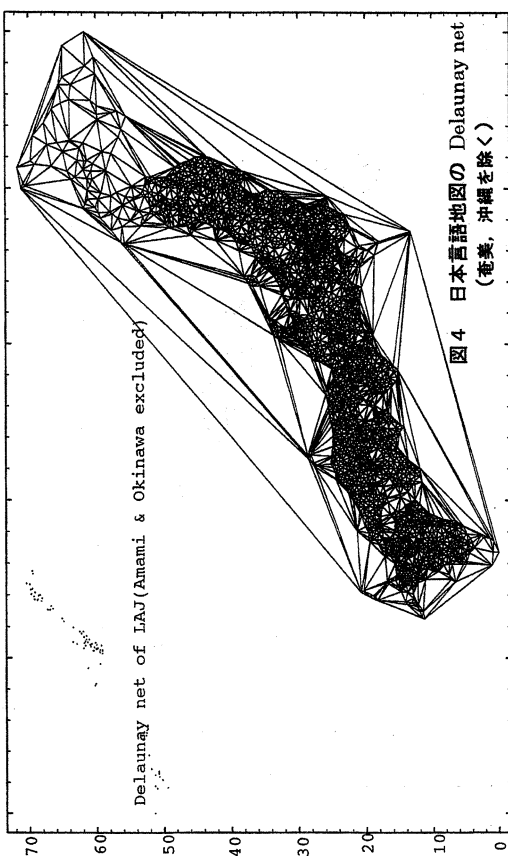


図4 日本語地図の Delaunay net
(奄美、沖縄を除く)

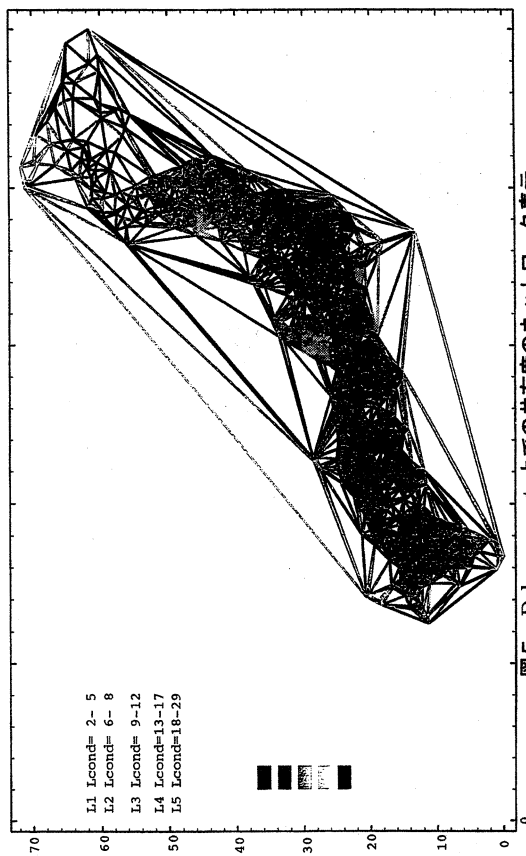


図5 Deluanay net 上での共有度のネットワーク表示
(カテゴライズ 5段階)

