

首都圏若年層の言語的地域差を把握するための方法と実践*

鎌水 兼貴

(国立国語研究所)

1. はじめに

東京を中心とする「首都圏」と呼ばれる地域は、人口の流動性が大きく、構成員も多様である。そのため、首都圏の言語は、地域の特徴がなく全国で通用するが、内部の細かい違いについては複雑だと考えられている。首都圏の言語の研究は、地域的に無標な「日本語」としておこなわれるか、もしくは言語意識や社会属性といった個人差の視点からおこなわれるのが一般的である。

首都圏は日本の政治・経済・文化の中心地であり、この地域の言語は、他地域に比べて共通語に強い影響を与えていることが予想される。現代日本語は、東京の言語を基盤として成立しており、東京を中心とする首都圏の言語動態の把握は、現代日本語の形成過程を理解する上で不可欠である。

しかし首都圏では、共通語と方言の差については意識されにくい。実際に首都圏の言語と共通語との差異はわずかであり、首都圏においては「方言」と呼ばれるような言語の地域差への関心は低く、言語地理学的調査も低調である。

首都圏の若年層であっても新たな言語現象は発生する。それは今後の日本語の形成にも重要な意味があると思われる。ところが、新しい言語現象が発見されても、首都圏において共時的にどのような状態で存在しているかについて、地理的調査を通じて把握しようとすることはあまりない。大都市は調査が困難な地域であるが、時期を逃すとすぐに言語使用の状態が変化してしまうおそれがある。

このような問題を解決するには、言語動態の把握にすぐれた探索的な調査手法の開発が課題となる。従来から、実施がしやすく、回答も短時間に大量に収集可能という点で、大学の授業における学生を対象としたアンケート調査が活用されてきた。しかし質問紙の集計には時間を要するため、調査結果を当該授業内で活用することが困難であった。授業内調査には、授業時間の妨害という側面があるため、1回の調査で良好な結果が出なかったからといって、繰り返し調査を実施することは容易ではない。

そのため、調査設計においてデータの効率的な収集が必要となる。インターネット経由、特に携帯電話の電子メールによるデータ収集は、個人識別可能な回答データが短時間で収集可能であ

* 鎌水兼貴(2013)「首都圏若年層の言語的地域差を把握するための方法と実践」(国立国語研究所論集 6, pp.217-243)からの転載。本稿は、国立国語研究所共同研究プロジェクト(萌芽・発掘型)「首都圏の言語の実態と動向に関する研究」(プロジェクトリーダー:三井はるみ <http://www.ninjal.ac.jp/shutoken/>)の研究成果である。また、2013年1月22日の国立国語研究所 NINJAL サロンでの発表を元に改訂したものである。

RMS システムの調査実験に協力していただいた各大学の学生の皆様に感謝の意を表す。また、国立国語研究所共同研究プロジェクトメンバーの皆様には RMS システムの利用面で多くの助言をいただいた。ここに御礼申し上げる。

り、結果報告までの時間を短縮することができる。そこで本研究では、携帯電話のメールによって収集した回答をリアルタイムで集計して言語地図の表示までおこなう、統合的な調査システム「RMS システム (Real-time Mobile Survey System)」を開発した。これは授業内調査の問題点を解決するだけでなく、調査結果を授業でも活用できるという利点がある。

本稿では、この新しい調査システムの概要を解説するとともに、実際に複数の大学の授業において実施した調査結果を示すことで、首都圏若年層に対して短期間で多人数調査を実施する意義について論じる。

2. 首都圏の言語について

2. 1 「首都圏」の範囲

本稿では、「首都圏の言語」という表現を用いるが、この「首都圏」の指す範囲はさまざまである。昭和31年に制定された「首都圏整備法」では、関東地方の1都6県（東京都、神奈川県、千葉県、埼玉県、茨城県、栃木県、群馬県）および山梨県を「首都圏」と定めている。しかし、内閣府（2011）では、「首都圏」を南関東の1都3県（東京都、神奈川県、千葉県、埼玉県）としており、根拠として国勢調査における「関東大都市圏」（総務省統計局2012a）の範囲が1都3県とほぼ同じ¹であることをあげている。

この「関東大都市圏」や、金本・徳岡（2002）の「都市雇用圏」などは、通勤・通学による移動を基準とした定義であり、ほぼ1都3県、都心からおよそ70kmまでの範囲を「地域的なまとまり」としてとらえている。2010年度の国勢調査における都道府県別昼夜間人口比率（総務省統計局2012b）をみると、東京都が118.4%と全国で最も高い反面、周囲3県は、神奈川県91.2%、千葉県89.5%、埼玉県88.6%と全国でも著しく低い。これは東京都に周辺3県からの通勤・通学者が流入していることをあらわしている²。

言語においても、「首都圏の言語」を「東京を中心とした言語的に均質と思われている地域的なまとまり」と考えるとき、地理的には1都3県に近い範囲とするのが妥当と思われる。

2. 2 首都圏若年層の言語について

首都圏の言語は、もともと共通語との類似性が高い。これは1950年代の全国規模の方言調査である『日本言語地図』（国立国語研究所編1966-74 以下、LAJ）の結果からもわかる。

図1は、1都3県における2つの時代の共通語使用率の調査結果である。左の河西（1981）は、LAJの82項目における都道府県別共通語使用率³であり、右の井上（1997）は、河西と同じ82項目の使用度を中学生に調査した「全国中学校言語使用調査」（以下、井上調査）の結果である。両調査は調査時の年齢が大きく異なっており、平均した生年でみると、高年層を対象とした

¹ 「関東大都市圏」は1都3県から埼玉県西部・千葉県南部を除き、栃木県南部・茨城県南部・山梨県東部を加えた地域であり、実質的に1都3県となっている。

² 東京都内でも100%を超えるのは、中心部の、ほぼ旧東京15区の範囲のみである。

³ 河西（1981）では「標準語」という用語を用いているが、「地図の見出し語」を便宜的に標準語とみなしているだけであるため、本稿では「共通語」とした。

LAJ はおよそ 1895 年生まれ、中学生を対象とした井上調査はおよそ 1981 年生まれとなる。調査時期の差は約 40 年だが、生年では 90 年近い差である。

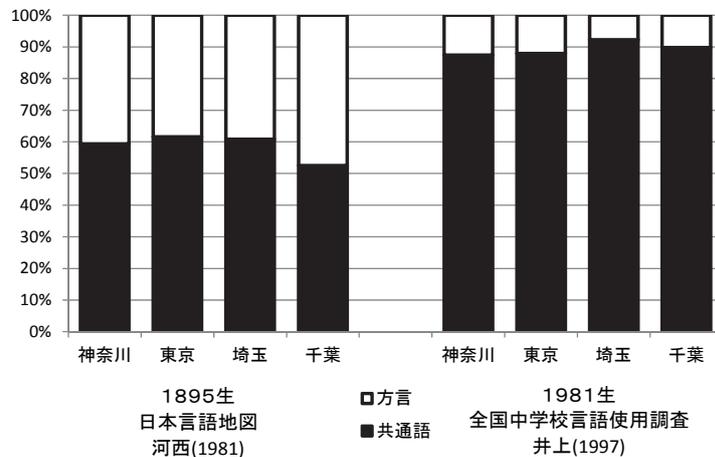


図 1 1 都 3 県における共通語形使用率

図 1 をみると、1 都 3 県の共通語形使用率は、LAJ の世代ですでに 6 割前後である。全国平均値である 37% に比べて非常に高く、もともと首都圏の伝統的方言が現代共通語と類似していたことがわかる。さらに井上調査の世代では 9 割が共通語形となっており⁴、首都圏の若年層にとっては「方言」というものを意識することが難しいことがわかる。

首都圏若年層におけるわずかな非共通語形についても、伝統的方言形の場合は、使用そのものが急速に衰退しているか（概念自体の衰退も含む）、使用される場合でも、使用意識において共通語に準じた扱いになっている（「俗語」「くだけた共通語」など）ことが多い。

首都圏では、戦後、移住によって人口が爆発的に増加した。移住者は首都圏にもとからある伝統的方言は採用しにくい。現在は移住 2 世、3 世の時代となっており、彼らは首都圏で生育した人々であるが、自分たちの使用する言語は、従来の伝統的方言とは異なる、地域性の希薄な言語だと意識する傾向がある。また、首都圏内部の地域的差異についても意識されることは少なく、言語的に均質だという意識があると思われる。

⁴ 実際にはこの 1 都 3 県のみならず、全国的に共通語使用率が高い。この調査で全国最低だった佐賀県でも 82.2% である。共通語能力の向上と方言能力の衰退は完全に対称ではないが、首都圏においては、もとの伝統方言形が共通語形と類似していることから、方言形が衰退しやすい環境にある。

2. 3 首都圏若年層における地理的調査の必要性

井上（1994）は「東京新方言」という術語を用いて、共通語使用の中核である首都圏の若年層であっても非共通語化現象が存在することを示している。さらに井上（2000）では「言語変化の雨傘モデル」によって、首都圏で使用される非共通語形が、「共通語」のように全国に急速に拡散することを示している。そして、その非共通語形の中には、首都圏周辺地域の方言形⁵の流入もあるとしている。ある方向から非共通語形が流入する場合、非共通語形が急速に拡散するまでは、かならず首都圏の内部にも言語の地域差が存在すると思われる。

しかし、1都3県の合計人口は2010年度の時点で3562万人（国勢調査 総務省統計局2012c）と巨大であり、「都市の多様性」という先入観が生じやすい。そのうえ1都3県は言語的に均質だという意識もあるため、たとえ言語的な差異がみつかったとしても、地域差の可能性は考慮せず、属性差（年齢・性別・学歴・所属集団など）や個人差（パーソナリティや意識）に目が向きやすいと思われる。

たしかに首都圏では、東京への通勤・通学者が多く、テレビや新聞などのマスコミも東京から発信されている。しかし首都圏の人々が共有しているのは中心部の東京だけであり、日常生活の場は、それぞれの生育地域にあると思われる⁶。多くの人は高校卒業までは生育地周辺で生活しており、言語形成における地理的要因は、他の要因と比べても決して小さくないと予想される。

以上から本研究では、首都圏の若年層において新規の言語現象が発生した場合、そこには地域差が含まれている可能性があると考ええる。そのためには、積極的に新規の言語現象を探索する必要がある。

若年層における地域差は、その形成過程を実時間で観察することができる。伝統方言のかつての地域差とあわせて分析することにより、首都圏の言語形成過程の解明につながるとと思われる。このことは現代日本語の形成過程にも関わる重要な研究だといえる。

ただ、首都圏のような大都市部では、調査の承諾が得にくい。もし首都圏の言語が地域的に均質であるなら、任意の一地点での調査だけで、属性差・個人差などを把握することができる。しかし地域差があるとすれば、多地点での地理的調査が不可欠である。首都圏で多くの調査対象者を探すことは容易ではない。調査地点数が少ない場合、たとえ差異を発見したとしても誤差の範囲であることを否定できない。そもそも探索的調査の場合は、多くの言語現象について地理的調査を実施しても、どれも首都圏全域で地域差がみられないこともありうる。

首都圏若年層の言語の地理的調査は、研究上の重要性は高いものの、労力が大きく、思わしい結果が出ない可能性がある。このような調査をおこなう場合、調査設計において相当の工夫が必要である。

⁵ 有名な例としては、ジャン（中部地方より）、ウザッタイ（東京都多摩地方より）、チガクナイ（北関東より）など。

⁶ 東京を中心とした放射状の交通手段が発達している反面、周辺地域間の移動については整備が遅れていることからわかる。たとえば埼玉県と神奈川県の人が行き来をする頻度は高くないであろう。

3. 首都圏若年層に対する効率的な調査

3. 1 調査コストを低くする必要性

前節で述べたように、首都圏の若年層は共通語使用率が非常に高く、地理的な言語動態をとらえようとしても、そもそもどのような言語現象に地域差があるのか、といったところから探さなければならぬ。

準備調査によって数多くの項目の調査を実施して、本調査の調査項目を選定できれば理想的だが、ある程度の人数と項目数を調査するだけでも一定のコストがかかる。アンケート調査であれば大量人数の回答を確保できるが、探索的調査である以上、簡単に望んだ結果が出るとは限らない。1回の調査であっても、調査票の準備から結果の集計まで、ある程度の時間を要する。しかも首都圏での言語現象は普及が早いため、あまり調査間隔を空けることができない。

非効率的な探索的調査を継続的に実施しようとするならば、1回あたりの調査コストを下げなければならない。そのためには、調査対象が限定されていても、研究者にとって実施しやすい環境の中で、繰り返し探索的調査を実施するのが現実的である。

多くの研究者にとって実施しやすい調査は、授業時に受講生に対して実施するアンケート調査だと考えられる。学生の生育地を用いて地理的分布をみる方法である。方法としては簡単にみえるが、大学の授業における調査の問題点について、あらかじめ検証する必要がある。

以下、大学生という対象の問題、学生の回答意欲の問題、授業妨害の問題の3つの面から考察する。

3. 2 授業時を利用した大学生に対するアンケート調査

授業における調査は、一種の集合調査である。受講生から短時間に効率よく回答を収集できる。大学における1つの授業は数か月にわたって開講されるため、同一学生に対する複数回の調査も可能な環境にある。つまり1回の調査が失敗したとしても、同一の調査対象に対して再調査が実施できるため、探索的調査に適しているといえる。

その一方で、大学生を調査対象とすることには、次のような問題点がある。

第一に、学校の立地による回答者の生育地の偏りの問題がある。大学の場合は高校のように立地周辺のみには偏ることはなく、大学を中心とした比較的広範囲のデータを取得することができる。ただし、1つの大学だけで首都圏全体をカバーすることは難しく、さらに広範囲のデータを収集したい場合には、複数の学校の授業で連携して調査を実施する必要がある。

第二に、学生の大学内での言語接触の影響についてである。たしかに大学進学後に普及した表現については、現在通学する大学の影響が強いことが予想される⁷。しかし大学生の年齢は、すでに柴田（1956）のいう「言語形成期」以降である。高校までに習得した表現については、ある程度固定化したとも考えられる。また大学生のほとんどが高校卒業後4年以内であり、遠方の大学に進学した学生も、移住先の言語の影響は限定的と考えられる。注意は必要だが、学生の生育地情報を用いて分布を調べることには意味があると思われる。

⁷ ここでは大学の立地する地域の言語の影響のことをさすが、「キャンパスことば」のような集団語については、地域ではなく所属する大学が差異に関係する。

第三に、大学生だけの調査は高学歴に偏っており、同世代の全体を代表していないという問題がある。そもそもサンプリング調査でない以上、統計的には世代を代表しているとはいえないが、それとは別に、高学歴層の回答は、規範的な方向に偏る可能性があり、特に俗語的扱いを受けやすい新しい表現については注意が必要である。しかし現在の日本の大学進学率は5割強と高く、大学生という社会的属性が少数派であるとはいえない。また、学術的な質問がしやすい点は、むしろ大学生であることの利点といえる。分析においても、調査対象の属性がある程度統一されることで解釈がしやすくなる面もある。

3. 3 回答意欲を高める調査

アンケート調査は、面接調査とは異なり、たとえ調査票内での指示を徹底したとしても、回答者側の全員が質問側の意図を正確に理解して回答しているとは限らない。質問文に疑問をもったままの回答や、質問の意図を誤解した回答があっても、その場で再回答を促すことができないため、どうしても一定の割合で無回答や無効回答、信頼性の低い回答など、不正確な回答が生じる。これに対しては、回答者数を増やして疑問回答を誤差の範囲に収めることや、逆にすべての回答を調べて疑問回答を排除する、といった処理が必要である。

田中（2004）は、アンケート調査における回答は、あくまで意識の反映であり、実態とは異なることがあると指摘している。そのため、結果に疑問が生じた場合には、回答だけでなく質問内容も疑わねばならない。しかし、こういった問題以前に、回答者が調査自体に不信感をもっていたり、調査内容に興味をもてない場合は、回答する意欲が低くなり、回答率が低下する可能性がある。

本研究の調査目的は、言語の地域差の探索であり、言語的背景は不可欠な質問項目である。しかし生年、性別、生育地、現住地、両親の出身地といった個人情報をたずねるため、回答を敬遠される可能性がある。回答してもらうためには、調査以前に個人情報を確実に慎重に扱うことが重要である。

授業内調査では、たとえ学生に調査の目的を説明したとしても、結果の提示が保証されない場合は、調査に対する興味が高まらない可能性がある。授業と関係した内容ではあっても、単に調査対象として学生を扱うのでは、学生側の好奇心を奪い、回答意欲の低下につながる。従来の質問紙調査では、データの入力作業、集計作業に多大な労力を要するため、学生に対して授業時間中はおろか、開講期間内に調査結果を示すことも難しい場合が少なくなかった。

学生の回答意欲を増すためにも、調査してから結果の出力までに時間がかからないことが理想的といえる。

3. 4 授業を妨害しない調査

調査者側の問題も考えなければならない。最大の問題は、授業中に調査をすることによる授業内容への影響である。授業時間内に調査を実施する場合、回答時間中の授業は中断せざるをえない。さらに調査結果をその授業内容に利用できない場合、実質的に授業の妨害となっていることを認識すべきである。他者の協力を得て、授業内調査を依頼する場合には、その授業に活用でき

ない調査を実施すること自体が問題となる可能性もある。

そもそも学生に対する調査は、教室における教員と学生の関係を利用した、半ば強制力をもった場における情報提供行為である。学生は、調査の内容にかかわらず、教員からの要請に応じなければならない状態におかれている。それだけに調査内容は、授業内容と関係があることが望ましい。授業内容に関係するのであれば、調査結果は迅速に授業で活用可能な状態にする必要があり、関係が低い場合には、最低限の時間で調査を完了できるようにしなければならない。

調査の所要時間は、授業の妨害時間でもある。つまり短時間で調査を完了することが、妨害を減らすことになる。もちろん質問数を少なくすれば調査時間が短くなるが、それでは調査の意味が薄れてしまう。質問数が多くても短時間で回答できるような方法を採用する必要がある。

集計結果の活用までを迅速にするためには、回答データの電子化までが短時間でおこなわれることが望ましい。質問紙のデータを手作業で入力するのではなく、WEB や電子メールなど、インターネット経由でデータを送信するなら、最初からデータは電子化されており、集計に要する時間は最小限となる。問題となるのは実施環境の整備である。PC 教室のような特殊な教室の使用を前提とする調査は実施可能な環境を狭める。通常の教室でも調査が実施できるような調査が望ましく、そのためには携帯端末等の利用が考えられる。この点については後で検討する。

調査の実施時間や、集計時間が短縮される場合、調査結果を踏まえて再調査を実施することが容易となる。すなわち、結果が思わしくない場合に質問を変えて再調査をしたり、学生の反応によって調査内容を追加したりすることができる。調査から結果公開までの時間を短縮することは、授業の妨害となる調査から授業に活用する調査へと変わることにつながる。このように、調査を授業中に活用するためには、結果に応じて質問の追加、変更を可能にする仕組みが必要である。

3. 5 授業で活用可能な調査システムの必要性

以上、首都圏若年層の地域差について、大学の授業を利用して探索的調査を実施する際の問題点を考察した。従来の調査方法とは異なる、新しい調査において必要な5点を、その効果とともにまとめる。

- | | |
|------------------------|--------------------|
| a. 同一回答者に対して繰り返し調査ができる | → 調査項目数を増やす |
| b. 他の授業での結果と組み合わせられる | → 回答者を増やす、対象地域を広げる |
| c. 迅速に結果を提示できる | → 調査が授業を妨害しないようにする |
| d. どの場所でも全員に実施できる | → 通常の教室での授業で活用する |
| e. 調査内容を臨機応変に設定できる | → 授業内容に沿った活用をする |

上記 a～e は相互に関係している。a と e によって柔軟な調査設計が可能となり、b と d によって大人数調査としての威力が発揮される。c はすべてに関係する。事前に準備した調査項目をたずねるだけでなく、学生から得た調査項目についての調査をその場で実施することも可能となるため、授業内容にも調査を組み込みやすくなる。また、他の授業のデータとも組み合わせる結果表示ができるため、他の授業で調査して良好な結果が得られた項目のみを調査する、といった使

用もできる。

目的はやや異なるが、木村（2009）は、高校などの学校への調査依頼時に考慮すべき点について、①サンプリングの問題、②学校に負担の少ない設計、③個人情報保護への配慮、④学校へのフィードバック、の4点をあげている。②の学校の負担については、コスト低減のために、研究者間でデータを共有する仕組みの必要性を提言している。これは上記 a～e の必要性にも通じるものがある。

本研究のような探索的調査において、調査の項目数と地理的範囲の両方を個人で広げることは容易でない。他の調査を参照でき、結果を組み合わせて使用できるようなシステムを構築することで、1回の調査は小規模であっても有効活用することができる。

集計結果を短時間に学生にフィードバックすることによって、授業内調査は「授業時間を妨害する存在」から「授業に活用できるツール」へと変化する。それだけではなく、学生と教員の関係も変化する。学生と教員が、従来の「調査を強制する側、される側という対立関係」から、「ともに授業を作り上げる協力関係」へと再構築されることになる。これは、学生の調査への参加意欲を向上させ、回答率の上昇や、無効回答の減少にもつながることが期待される。

アンケート調査を授業に活用する試みは、これまでは個人の努力によって実現されてきた。授業で調査を実施し、次回の授業までに集計結果や言語地図を示すのであれば、特別なスキルがなくとも、ある程度は実現可能である。しかし調査者個人の労力に大きく依存する場合、現実的には誰もが可能とはいえない。前述の a～e を満たすような調査を、特別な知識を必要とせず、誰でも実施できるようにするためには、技術的な補助が不可欠であり、専用の調査システムの開発が必要となる。

4. 言語地図形式による回答結果の自動出力

4. 1 言語地図作成の工程

新しい調査システムの開発においてデータ収集とともに重要なのが結果提示の迅速化である。特に結果提示において、言語の地域差を示すためには言語地図が必要となる。ここでは言語地図作成の自動化について、先行研究から考察する。

言語地理学的研究において言語地図は分析の中心をなすものである。調査の規模にかかわらず、調査終了から言語地図の完成までには、以下の工程を経る。

A 回答の電子化 → B 回答の整理 → C 地図化・集計処理

A の回答の電子化は、調査者が調査票に記入した回答内容をデータ化する作業である。電子化以前はカードへの転記であったが、現在では、電子データで処理することが一般的である。大規模調査になるとデータ入力作業は多大な労力となる。選択式回答の場合にはマークシート方式や、WEB アンケートのような電子入力方式が効率的であるが、専用のシステムが必要となる。

B の回答の整理は、選択式回答の場合を除いて自動化は困難である。自由記述式回答の整理作業は単純ではなく、整理方法に分析的観点を含むため、C の地図化・集計処理の結果を受けて、

繰り返し回答の整理をおこなう必要がある。自動化は困難であるが、回答の異なり数が多い場合には、回答データの整列、検索等、コンピュータを利用して整理を補助する処理をおこなう。通常は、表計算ソフトなどを用いて人がおこなう。

Cの地図化・集計処理については、手作業によって作成していた時代は、最も労力を要する工程であった。集計処理に関しては、単純集計であれば手作業で回答の書かれたカードを分類しなければならなかった。しかしコンピュータの普及以降は、回答データを電子化することにより、個人でも表計算ソフトや、統計ソフトを利用することで、処理が可能になった。一方、地図化作業についても、白地図上に回答に対応した記号を繰り返し押印する作業が続くため、大規模な調査ほど言語地図の完成までに時間を要していた。作業工程が複雑であるため、コンピュータが普及してからも個人で言語地図を作成することは難しかった。そのため言語地図作成の高速化、省力化は、さまざまな研究者によって試みられてきた。

以上から、調査結果を迅速に提示するためには、Aの回答の電子化部分と、Cの言語地図の作成部分を高速化する必要がある。以下、それぞれについて検討する。

4. 2 言語地図作成の自動化

まず、Cの言語地図の作成部分について概観する。言語地図の電子化の歴史については岸江(2007)によってまとめられている。基本的な作成原理は大型計算機を用いた徳川・山本(1967)によってすでに確立されていた。1950年代に国立国語研究所で調査された『日本言語地図』のための膨大な調査データの整理を効率化する目的で大型計算機による言語地図の作成が検討された。地点番号(緯度、経度に対応)を座標として言語地図をプロットする手法である。グラフの散布図と同等の手法で実現できるため、非常に単純である。地図の出力についても、荻野(1975)においてX-Yプロッターを利用することで、高品質の地図を実現している。荻野(1981)の開発した言語地図作成システムGLAPSは、こうした言語地図作成ツールの先駆けといえる。

1980年代になると、パーソナル・コンピュータの普及により、最終出力だけでなく、画面上でもグラフィカルな地図を扱うことが可能になった。これにより、福嶋(1983)のSEALや、前川(1988)のEGLなど、使用者に配慮した、対話型の言語地図作成ツールが開発された。大西(2002)は、『方言文法全国地図』の編集作業において、Adobe Illustrator用のプラグインとして言語地図を作成するシステムLMSを開発し、研究用だけでなく、大型版の版下としても耐えうる高品質な言語地図を作成できるようになった。近年では、地理学における地理情報システム(GIS)の普及により、単に地図を作成するだけでなく、統計処理と組み合わせた複雑な解析も可能になった。これにより分析方法も発展するものと思われる。

しかし、これらのツールを使用するためには、ある程度の技術の習得が必要であり、誰でも使用可能なものとしては認知されていない。コンピュータ上で言語地図を作成するための一般的な解説書は、GISソフトであるMANDARAやArcGISの利用について書かれた中井(2005)しかなく、現在でも、言語地図の作成には相当のハードルがあることがわかる。

4. 3 回答データ入力自動化

つづいて A の回答の電子化部分について述べる。面接調査の場合は、調査者が調査票に手書きで記入するため、現在でも電子データ化は手作業による入力が一般的である。ここではアンケート式の調査結果の電子化について、速報性の観点から概観する。

回答データの電子化において、最も効率的な方法は、回答者が回答時点で電子データとして入力することであろう。しかしインターネット登場以前は、回答者と調査者の間のデータのやりとりが問題であった。フロッピーディスク等の物理的な記憶媒体のやりとりには手間がかかる。ワークステーションの端末に全員で接続することによって共有することも可能であったが、専用の端末室での作業が必要となる。

1990年代後半よりインターネットが普及すると、データ収集部分が電子化された研究例がみられるようになった。田中（2004）や荻野（2004）によって電子メールを利用した調査手法が紹介されたほか、中井（2005）は、電子メールを用いたアンケート結果を GIS ソフトで地図化する方法について解説している。大西他（2011）による「方言メール調査」システムでは、電子メールで回答を収集し、自動的にデータベース化をおこなっている⁸。

選択式回答の場合、電子データの自動収集が可能になると、言語地図作成の全工程を自動化する研究もあらわれた。林・日高（2008）は WEB を用いたアンケートシステムを開発し、データ収集から地図表示までを自動化した。PC 教室での授業を想定して開発されており、授業時間中に調査から結果提示まで可能であり、学生にとっても調査の意義を理解しやすい。

このほか、1990年代に開発された高橋（2003）の SUGDAS は、データ入力部分の自動化はされていないものの、授業内調査と短時間での報告を念頭において、電子データの入力から言語地図の出力までの流れを見通している点で、本研究の趣旨に近い。SUGDAS は選択式回答が中心であったが、高橋（2007）は『方言文法全国地図』の電子データを分析するために改良した GAI-Sugdas で、回答語形を画面上で選択する方式によって、回答の整理の省力化も図っている。電子データの利用を前提としたシステムとして非常に重要な位置にあるといえる。

5. 携帯メールを用いた調査

5. 1 インターネット経由での回答データの収集

前節で述べたように、調査結果を迅速に提示するためには、データ収集段階でインターネットを使用することが必要である。回答者が電子メールや WEB 経由でデータを送信することにより、従来の質問紙による調査で必要だったデータ入力作業が不要となり、データ利用までの時間が短縮できる。ただし PC や携帯端末などの電子機器の利用度には個人差がある。インターネットの利用度も社会的属性の 1 つであり、橋元（2004）はインターネットをよく利用する人の回答の偏りに言及している。しかし授業内調査の場合は、教室内の全員が対象となるため、この問題は避けられる。

荻野（2004）は、自身の実施した電子メール調査の経験をもとに、PC における、電子メール

⁸ データの自動収集部分は筆者が作成を担当しており、本研究の RMS システムの元にもなっている。

とWEBの調査の比較をおこなっている。電子メールの長所については、(1) 回答者にとって調査の全体がわかる、(2) 回答者の心理として返信しやすい、(3) 回答者がアドレスによって特定できる、(4) 回答者がどういう集団であるか把握できる、という点をあげている。

一方、WEBの長所については、(1) データが形式化されている、(2) サブクエスチョンへの対応が容易である、(3) メール調査よりも利用者が回答しやすい⁹、という点をあげている。

電子メールでは回答者の特定が、WEBでは調査の流れのコントロールが、それぞれ利点とされていることがわかる。荻野(2004)は結論として、連絡を電子メールでおこない、メール内に示したアドレス経由でWEB調査を実施することを提唱している。現在多くのWEBサービスが、登録にメールアドレスを採用していることから、妥当な方法であるといえる。

荻野の研究から10年近くが経過し、若年層の間では、新しいコミュニケーション・ツールの利用が増加した。現在主流であるTwitterやFacebook、LINEといったコミュニケーション・ツールを調査に活用することも考えられ、実際に多くの調査で利用されている。しかし、こうしたツールは企業のWEBサービスの延長であるため、会員登録していない人は利用できない。授業内調査は受講生全員の参加を目指すため、古くから存在する、ほぼ全員が利用できる電子メールの利用が適当と思われる。

つづいて授業場面でのデータ収集方法について検討するため、アンケート調査の媒体別特徴を表1にまとめた。電子メールについては携帯電話のほかにPCやWEBでの利用があるが、重複点が多いので省略した。以下、WEBを用いた調査と、携帯電話の電子メール(以下、携帯メール)を用いた調査について、質問数、調査場所、料金的負担、速報性、複数回調査という調査コストに関わる観点から検討する。

表1 アンケート調査の媒体別特徴

調査媒体	紙	WEB	携帯メール
質問数	○ 回答によって質問が分岐するような場合には、わかりにくくなる。	○ 調査の流れをプログラムでコントロールできる。	× 最大1通20問が限度。複雑な質問は難しい。
調査場所	○ どこでも可能。ただしその場で実施しない場合、回収率が下がるおそれがある。	△ WEB利用環境が必要。携帯からのWEB利用であればどこでも可能。	○ 携帯メールが使用できれば、どこでも可能。
料金的負担	○ 全くかからない。逆に調査者側に印刷費や郵送費などの負担が発生する。	△ PC教室で行う場合は負担がない。携帯でのWEB利用は人によっては料金的負担が大きくなる。	△ 携帯メールは日常的に使用する人が多く、料金的負担が気にならない人が多い。
速報性	× データ入力に時間がかかる。	○ 最初から電子データとしてデータが得られるため、サーバー側のプログラムを用いることで自動集計が可能。	○
複数回調査	△ 回答者が前回と同一人物であることを確かめるため、IDの発行等が必要となる。大学であれば学籍番号が使いやすい。	△	○ 携帯電話は個人と結び付いているため。メールアドレスで同一人物か判断可能。

⁹ WEBの長所について荻野自身の意見は(1)(2)だけである。(3)は調査時の回答者の意見として紹介している。

5. 2 WEB 調査と携帯メール調査の比較

表 1 をみると、WEB による調査は欠点が少ない。質問紙に近い複雑な調査も可能であり、回答欄の整合性についても、プログラムを作成すれば自動的にチェックが可能である。入力ミスが減少し、効率的であるため、調査会社で広く使用されている方法である。

大学の授業で WEB による調査を実施する場合、PC 教室においては容易に実施可能である。すでに前述の林・日高（2008）において実現されている。PC のない通常の教室においては、学生の所有する端末、すなわち携帯電話を利用することが考えられる¹⁰。しかし、携帯電話の利用には通信料金が発生する。スマートフォンの急速な普及により、携帯電話からの WEB 利用も増加しているが、全員が利用しているわけではない。携帯電話で WEB を利用しない学生にとって、WEB の利用は料金的負担が大きい¹¹。このため、全学生を対象とする場合、携帯電話からの WEB 利用には慎重であることが望ましい。

一方、携帯メールの場合、ほぼ全員が携帯電話を常時所持しており、所有者以外の利用が基本的にない点が特徴的である。携帯端末と結合したメールアドレスから回答者を特定できるため、同一人物に対して複数回調査をおこなう場合に、個人 ID の発行やログイン処理などをすることなく、データの蓄積が可能である。これは回答側の入力負担の軽減にもつながる。その反面、メールアドレスを蓄積することになるため、メールアドレスの流出防止等、調査システムの開発において注意を払わなければならない。

また、WEB よりは安いものの、携帯メールにも料金負担は存在する。携帯電話所持者のほぼ全員がメール機能を利用していると考えられるため、調査における費用負担感は低いと思われるが、負担が発生している点は意識する必要がある。

電子メールの入力欄には自由に内容を入力できてしまうため、回答欄の制御は不可能である。質問数が多いと入力ミスを誘発するおそれがある。荻野（2004）の調査では、調査票自体が電子メールの本文に入っており、指定の場所に追記して返信してもらう、という仕組みであった。携帯電話でも同様の返信はできるが、携帯電話の小さい画面では、長文の調査票を表示すると、改行場所の違いなどにより表示が乱れることで、返信が困難になる可能性がある。携帯メールによる調査は、調査規模に制約があるといえるだろう。

以上から、携帯メールは、少ない質問を繰り返し実施する、本研究のような調査に適しているといえる。田中（2006）は、「新規登場事象」の探索的な調査手法として携帯メールの活用を提案している。現代の若年層は携帯電話での文字入力に慣れており、ほぼ全員が短時間に入力することができる。頻繁に調査を実施しても支障はないと思われる。これまでは学生に手を上げさせていたような、授業中に気軽に質問するような調査も、携帯メールを使ったシステムであれば実現できる可能性がある。

¹⁰ 学生全員に iPad のような通信端末を配布する大学もある。この場合は、通常の教室においても全員に WEB 調査が可能であるが、一般的ではないため、通常は個人所有の端末を利用する必要がある。

¹¹ 学内に、学生用の無線 LAN が整備されていれば、学生への料金的負担はない。

6. RMS システム

6. 1 RMS システム概要

以上から、首都圏若年層の言語の地域差を探索するためには、授業中に学生が携帯メールで回答し、その場で調査結果である言語地図が出力できるようなシステムが、最も効果的だと思われる。そのため本研究では専用の新しい調査システムを開発した。この調査は、「リアルタイムに結果を見ることができる携帯電話を利用した調査」であることから、“Real-time Mobile Survey (RMS)” (リアルタイム携帯調査) と命名し、RMS を実施するために開発した調査システムを「RMS システム」と呼ぶことにする。

RMS システムは、①調査者が調査項目を登録する、②回答者が調査項目に対して回答のメールを送信する、③調査結果である言語地図が WEB 上に表示される、という 3 段階の流れから構成される。データは自動的にデータベースに蓄積され、①②③のプログラムからアクセスされる。

図 2 に RMS システムの概要を示す。システムは、インターネットに接続した専用サーバーを利用することを前提とし、以下の 4 部分に分かれる。

- ①調査管理 (調査票や調査者情報を管理する)
- ②回答処理 (電子メールのデータを解析する)
- ③結果表示 (回答の整理や地図表示をおこなう)
- ④データベース (データの格納をする)

RMS システムでは、④のデータベースは、専用の管理システムは利用せず、単にテキストファイルをフォルダに分類して格納している。①②③の部分は、プログラムによって実現される。

現在開発している RMS システムは、サーバーはレンタルの UNIX サーバー、プログラムは Perl スクリプトを利用している。Perl スクリプトは筆者が作成した。

なお、調査票について、現在のシステムでは、回答者に質問文を見せるための機能が考慮されていない。調査者が別途作成した質問用紙を配布するか、もしくは黒板 (ホワイトボード) やプロジェクターで提示することを前提としている。しかし、言語調査の中には、絵を見せたり、音声を聞かせたりする質問項目もある。将来的には③結果表示の部分を、結果だけではなく、質問文の提示にも使用できるように変更する必要があるだろう。総合的なプレゼンテーションをする部分として発展させる予定である。

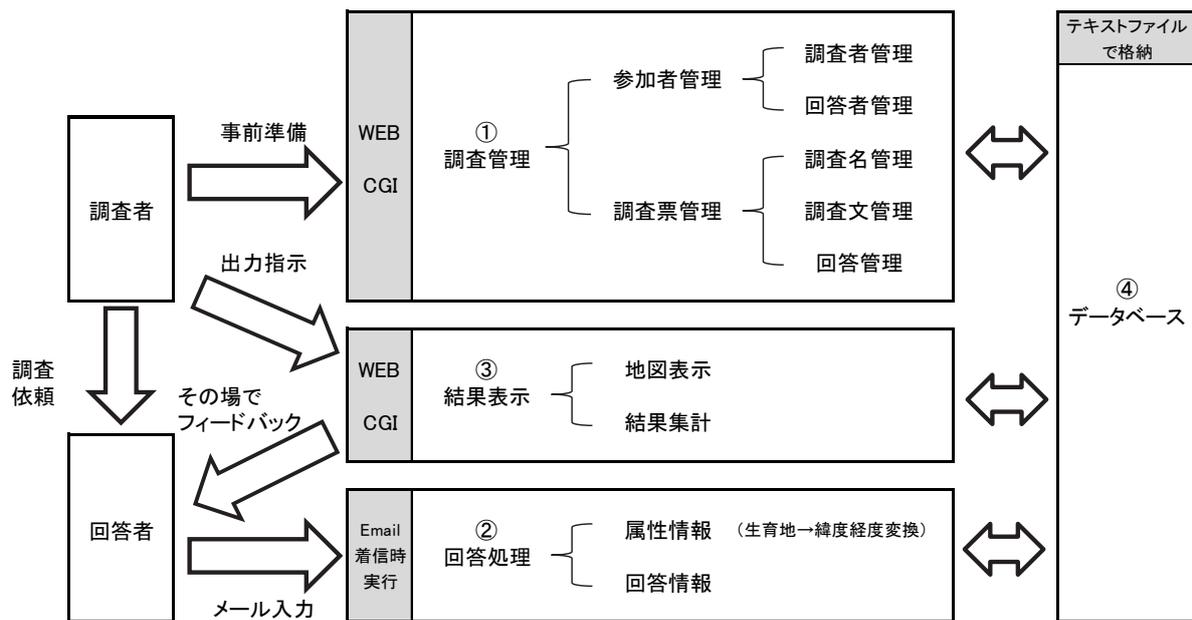


図 2 RMS システム全体図

6. 2 各段階の説明

図2で示したRMS システムのそれぞれの部分について解説する。調査システムは鑑水(2011a)や鑑水(2012)でも解説がなされているが、その後若干の改良が加えられている。執筆時点(2013年7月)でも未実装の機能が存在するが、今後の課題として解説する。

6. 2. 1 調査管理

図2の①調査管理は、実施する調査そのものを管理する部分である。Perl スクリプトであるrms.cgiによって実現される。インターネットからのアクセスは、たとえば、

<http://xxxx.xxx/rms.cgi> (xxxx.xxx は実装するサイトのアドレス) のようになる。内部は、(a) 参加者管理と (b) 調査票管理という2つの処理からなっている。

(a) 参加者管理

調査者、回答者はともに調査の参加者である。さまざまな参加者による調査を一括して管理するため、参加者の管理は重要な部分となる。参加者管理は、(1) 調査者管理と (2) 回答者管理の2つに分かれる。

(1) 調査者管理： 個人的な調査をする場合と、複数の調査者でグループを作る場合がある。調査者には、個人の管理のほかに、グループ単位の管理が必要となる。システムへのアクセス権の管理もするため、調査者はIDとパスワードを登録する。これはrms.cgiと、後述する結果出力のためのmap.cgiにアクセスする際の認証に用いられる。

(2) **回答者管理**： 回答処理中の属性情報の登録によっておこなわれる。複数の授業に出席する学生によって回答者が重複することがある場合には、回答の共有レベルの指定に従って、回答者を統合することができる。現在は実装されていない。

(b) **調査票管理**

調査は調査票単位で管理される。しかし、調査と調査票は一対一で対応しない。同一の調査票で異なる対象に実施する場合もあれば、異なる調査票を同一の対象に実施する場合もある。調査票自体が改訂されることもある。

異なる調査結果同士を比較しやすくするために、システム上では、個々の質問文単位で管理している。それぞれの質問文には、調査者が管理しやすいうようにタグを付与することができる。これにより結果出力時にタグを参照して、複数の調査の結果を組み合わせることができる。高密度地図の作成において有用である。

調査票管理は、(1) 調査名管理、(2) 調査文管理、(3) 回答管理の3つに分かれる。

(1) **調査名管理**： 複数の調査文をグループ化し、1つの調査として扱うことができるようにする。現在は実装されていない機能であり、今後の課題である。調査文はタグ付けや階層化などにより整理する必要がある。「調査項目データベース」のような独立したデータベースを別に作成して、RMS システムから参照する方法も考えられる。

(2) **調査文管理**： 質問文、質問の種類、質問のタグ、絵や音声の有無、選択肢、関連質問など、調査文に関する情報が管理される。作成者は調査文ごとに共有許可を与えることができる。

(3) **回答管理**： 回答情報を「回答者×質問項目」という1枚の表データとして扱う。その表中の1セルを選択することで、回答に関する詳細情報の閲覧ができ、修正・整理をおこなうことができる。表は、属性情報と回答情報とにわけて表示される。修正画面では、元のメール内容も表示され、データの変換が正しいかをチェックすることができる。回答ミスによる項目のずれの修正などもここでおこない、自動的にデータベースに結果が反映される。回答データのダウンロードや、外部データのアップロードもここでおこなう。そのため質問紙でおこなった調査についても、外部データとしてアップロードすることで、RMS システムで利用することができる。

6. 2. 2 **回答処理**

図2の②回答処理は、調査実施時に回答の電子メールを受信する部分であり、システム全体の重要な部分の1つである。サーバーが、指定のメールアドレス宛のメールを受信した際に、メール転送機能を用いて、メール内容がPerl スクリプトに送られる。Perl スクリプトは、メール内容を解析し、回答者の属性情報、質問への回答情報にわけて、データベースに格納する。そのため、RMS をレンタルサーバーで実現する場合、メール転送時のスクリプトの実行が可能であるか確認する必要がある。もし不可能な場合には、一定時間ごとにスクリプトを実行するような機能

(cron) が必要となる。

回答処理は、(a) 属性情報と (b) 回答情報にわかれる。RMS では、質問番号がアルファベットの場合に属性情報、数字の場合に回答情報とみなされる。

(a) 属性情報

RMS 調査において、回答者は、準備なしに調査に参加できるわけではない。あらかじめ属性情報を送信することによって、メールアドレスが ID となり登録される。これにより再度調査をする場合でも属性情報の登録が不要となる。

件名に①調査管理で指定した「調査名」を指定することで、自動的にメールはその調査名の調査に参加する人の属性情報とみなされる。町丁目単位の生育地、転居歴、性別、配偶者・両親の出身地など、あらかじめアルファベット記号と属性情報の関係を調査文管理で指定しておくことによりデータベースに登録される。

図 3 は、属性情報のメールの例である。RMS 調査において位置情報の詳細度は、個人情報と研究のバランスを考慮して、町丁目単位としている。a,b,c は回答者の生育地に関する質問であるが、あらかじめ調査文管理で a を都道府県、b を市区町村、c を町丁目と指定しておくこと、この電子メールをサーバーが受信する際に、東京大学空間情報科学研究センターの「シンプルジオコーディング実験」を利用して自動的に緯度経度に変換される (図 4)。こうしてメールアドレスと位置情報が結び付けられる。

なお、メールアドレスと生育地などの個人情報が一覧となって収集されるため、個人情報の取り扱いに十分注意する必要がある。ただし RMS システムでは、メールアドレスには自動的に ID が振られ、画面上では ID のみが表示される。そのため調査者が直接メールアドレスを閲覧することはできないようになっている。

調査時に携帯電話を忘れてしまうことや、電池の残量不足で携帯メールが使用できない場合がある。こうした場合に対処するために、登録時に特定の ID を 1 つ以上登録しておくことが推奨される (学生対象の場合には学籍番号や氏名が一般的)。これにより、WEB から入力した場合でも回答の蓄積が可能となる。現在 WEB 入力については未実装である。

メール作成 (新規)

宛先	ans@chosa.in
件名	kaito
本文	a 東京都 b 立川市 c 緑町 d 男 e 1993 f 東京都立川市

図 3 属性情報のメール例

生育地	緯度	経度	信頼度
東京都江戸川区平井	35.703	139.85	5
静岡県静岡市清水区興津中町	35.056	138.52	5
兵庫県相生市赤坂	34.814	134.48	5
東京都新宿区東五軒町	35.706	139.73	5
埼玉県上尾市小泉	35.969	139.57	5
神奈川県厚木市松枝	35.448	139.36	5

図 4 住所から緯度・経度への変換例

問題点としては、前述したように、メールアドレスや学籍番号、氏名といった個人情報が自動的にデータベース化されることがあげられる。システム上、調査者がメールアドレスを閲覧できないようにしているとはいえ、個人を特定可能な情報に関しては、特に学生の許諾を得ていない場合、一定の調査期間が終了したら、データベースから削除することが望ましい。将来的には自動的に切り離す機能が必要であろう。ただしこの場合、同一人物が複数の調査で回答した場合、別の人として扱われる¹²。

(b) 回答情報

回答データも、件名に調査管理で指定した「調査名」を指定することで、自動的にその調査に対する回答とみなされる。本文には、行頭に設問記号である数字が入り、つづいて選択肢や自由回答が入る。

図5は回答情報のメールの例である。問1～3は選択式で、問4・5は記述式となっている。選択式の場合は回答の整理の必要がないため、すぐに地図出力が可能である。

また、調査名を空欄にすることで調査者が事前に登録をしていない状態でも調査をおこなうことができる。回答の制限時間を設ける場合には、設問記号も不要となり、本文に回答するだけでデータの蓄積が可能である。このような臨時的な調査については現在は未実装の機能である。

データの登録状況を後から確認可能にするため、送られてきたメールは、①オリジナルのメール、②調査ごとに振り分けたメール内容を指定方法によって変換したデータ、という2つの状態で蓄積される。このとき調査名が不明のメール(件名にミスがある場合)は、調査名不明(unknown)として蓄積される。②の指定方法による変換とは、あらかじめ文字コード・改行コードの変換や、表記(半角・全角、ひらがな・カタカナなど)の統合のような、異なるメール環境から送信されても正常に処理できるように基本的な処理をおこなうことである。通常はこの処理で問題はおきかないが、文字化け等や設問の区切りミスといった問題が発生した場合には、回答管理において修正作業をおこなうことができる。②にミスがないことを確認したら、①のオリジナルメールは破棄することができる。

メール作成(新規)	
宛先	ans@chosain
件名	kaito
本文	1 a 2 b 3 a ; b 4 ヤナサッテ 5 かわいい感じ

図5 回答情報のメール例

¹² 現在は同一期間に異なる授業でRMS調査が実施される場合、同一の学生が何度もフェイス情報を入力することになる。このためメールアドレスの共有条件についても検討する必要がある。

6. 2. 3 結果表示

図 2 の③結果表示は、集まったデータの処理に関する部分である。Perl スクリプトである、map.cgi によって実現される。インターネットからのアクセスは、たとえば、

http://xxxx.xxx/map.cgi (xxxx.xxx は実装するサイトのアドレス) のようになる。内部は、(a) 地図表示、(b) 結果集計の 2 つにわかれる。

(a) 地図表示

結果の出力部分は、回答者へのフィードバックとなるものであり、RMS システムにおける最も重要な部分の 1 つといえる。データベースに蓄積されている回答データを、言語地図という形で視覚的に表示する部分である。

まず実施した調査名を選択し、それから調査項目を選択すると言語地図が表示される。言語地図は、緯度経度の情報をそのまま座標に置き換えて表示する、最も単純な正距円筒図法を採用している。日本地図を作成する上で、岡本義雄氏（大阪教育大学）作成の「日本列島海岸線データ & 県境データ」を加工して利用している。国土地理院の「地球地図日本」のデータも利用可能である。

言語地図は、質問文ごとに表示されるが、そのときに前述の調査文管理で付与されたタグと同じタグをもつ調査の一覧が表示される。その一覧を選択することで、複数の調査結果を重ね合わせた地図が自動的に出力される。同一の質問であっても、参加者管理で異なるグループとして登録されている調査者の調査は一覧に表示されないため、調査データの無断利用を防止することができる。

言語地図は PNG 形式の画像として出力され、利用することができる。現時点では、言語地図しか出力できないが、地域差がない項目については、言語地図よりもむしろグラフ表示のほうが効果的である。そのため今後は棒グラフや円グラフ、折れ線グラフといった表示についても実装する予定である。

(b) 結果集計

結果集計は、回答の基礎集計出力をおこなう部分である。主に自由記述式の回答に利用する部分であるが、現在は回答の一覧と回答者数、割合しか出力できない。

選択式回答の場合は、ほぼ自動的に集計することが可能だが、自由記述式の場合は、回答の整理に時間を要する。そのため、高橋（2007）の GAJ-Sugdass にあるような、質問ごとの異なり語形を出力し、回答の整理を補助する機能を実装予定である。回答が多岐にわたる場合には、異なり語形を出力するだけでは分類の困難さはあまり改善しない。そこで鎌水（2007）の回答語形間の類似度を用いた自動分類などを利用して、効率性を高めていきたいと考えている。

6. 2. 4 データベース

①調査管理や②回答処理で蓄積されたデータは、図2の④データベースに格納される。③結果表示をおこなう際には、このデータを利用する。データベースは、テキストファイルとして、調査者単位でディレクトリにて管理される単純なものである。調査人数が大量になった場合は、データ処理に時間を要することが予想される。将来的にはデータベース管理システム (DBMS) を利用するように修正することも考えられる。

個々のデータファイルの書式は独自のものであるため、現状では利用しやすい形式に変換しなければならない。今後は鎌水他 (2010) が提案したような、XML を用いた方言調査に適したデータ記述方式に対応することも検討する必要があるだろう。鎌水 (2011b) は、異なる方言資料を統一的に扱うことを提案しており、その際に扱いやすい形式であることも重要だと思われる。

フェイスシートのデータには回答者の個人情報が含まれるため、データの漏洩を防止しなければならない。メールアドレスは、②回答処理で ID 番号に変換して、回答データとは別のデータとして格納している。しかし今後は暗号化など、セキュリティ面を考慮した処理についても検討する必要がある。非常に重要な課題である。

6. 3 複数の調査の組み合わせ

以上が RMS システムの概要である。最後に、複数の調査を組み合わせた場合のデータの扱いについて述べる。表2は、複数の大学、複数の授業における架空の調査例である。4日間で以下の調査がおこなわれたとする。

- (1) 火曜日の2時間目に、A大学の授業①で、事前に準備した調査1(6問)と、授業中に思いついた新たな調査2(2問)を実施した。
- (2) 火曜日の3時間目に、A大学の授業②で、事前に準備していた2問のほかに、前の時間の授業①で良好な結果が出た調査1の3問と調査2の1問を加えた調査3(6問)を実施した。そして授業中に新たな調査4(1問)を実施した。
- (3) 2日後の木曜日の2時間目に、別のB大学の授業で、A大学で良好な結果が得られた質問からなる調査5(5問)を実施し、授業中に新たな調査6(1問)を実施した。
- (4) 最初の調査から一週間が経過した火曜の2時間目に、A大学の授業①で、他の授業で良好な結果が得られた質問からなる調査7(2問)を実施した。

以上の調査の、質問文と回答者の関係を表示したものが、図6のイメージである。

表2 複数授業でのRMS 調査実施例（数字は質問数）

授業時間	授業名	調査	事前	他授業	授業内
火曜2限	A大学授業①	調査1	6		
		調査2			2
火曜3限	A大学授業②	調査3	2	4	
		調査4			1
木曜2限	B大学授業	調査5		5	
		調査6			1
火曜2限	A大学授業①	調査7		2	

		質問1	質問2	質問3	質問4	質問5	質問6	質問7	質問8	質問9	質問10	質問11	質問12				
A 大学 授業 ①	回答者1	○	○	○	△	△	△	×	×	×				← 調査7			
	回答者2	○	○	○	○	△	△	×	×	×	○	○	△		△		
	回答者3	○	○	×	×	×	△	△	△	×	×	×	○		○		
	回答者4	○	○	×	×	×	△	△	○	○	○	△	△		×	×	
	回答者5	○	○	×	×	×	△	△	×	×	×	○	○		△	△	
	回答者6	○	○	×	×	×	△	△	△	×	×	×	○		○	○	
	回答者7	○	○	×	×	×	○	○	○	△	△	×	×		×	×	
	回答者8	×	×	×	△	△	△	△	×	×	×	○	○		△	△	
	回答者9	○	○	×	×	×	△	△	△	×	×	×	○		○	○	
授A 業大 ②学	回答者10							○	○	○	△	△	△	×	×	×	← 調査4
	回答者11							×	×	×	○	○	△	△	△	○	
	回答者12							△	△	△	×	×	×	○	○	○	
	回答者13							△	△	△	×	×	×	○	○	○	
B 大 学 授 業	回答者14							×	×	×	○	○	△	△	△	○	← 調査6
	回答者15							△	△	△	×	×	×	○	○	○	
	回答者16							○	○	△	△	×	×	×	○	○	
	回答者17							×	×	×	○	○	△	△	△	○	
回答者18							△	△	△	×	×	×	○	○	○	○	
人数	8	9	9	9	18	18	18	9	18	18	4	4	14				

図6 複数の調査回答の組み合わせイメージ

図6の表の下に示されている数字が、質問ごとの総人数である。良好な結果が得られ、複数の授業で追加調査がされた場合、多くの人数の回答が得られることになる。

このとき、前述の調査文管理によって、同一質問に同じタグを付与することによって、これらのデータは統合されて自動的に出力することが可能となる。

一回の調査では少人数の回答しか得られないとしても、複数の授業で多人数の回答を自動的に表示することができるのが、RMS システムの特徴といえる。

7. RMS システムを利用した調査例

7. 1 関東方言形カタス

RMS システムを利用した首都圏若年層の調査の例として、方言形カタスの分布についてとり上げる。

「片付ける」の方言形であるカタスは、北関東を中心に関東地方の広範囲で使用されてきた。井上（1985）の東京～福島間のグロットグラムの結果や、井上編（1988）の東京都と神奈川県に

おける2世代言語地図から、神奈川県で使用されていないことや、若年層ではやや衰退傾向にあることなどが判明していた。しかし早野（1996）は千葉県松戸市の調査から、移住者や二世がカタスに東京的イメージをもったために、再び普及傾向に転じたとしている。

しかし、これらの先行研究からは、関東地方におけるカタスの普及状況を把握することができない。そこで大学生の関東地方での使用分布についてRMS調査を実施することにした。

7.2 カタスのRMS調査

カタスは、RMSシステムの調査実験開始当初（2011年6月）から、調査項目に採用され、ほとんどの調査実験で調査されてきた。そのため2013年6月までに1800人余りの大学生の回答を得ている。

それぞれの授業の調査での質問番号は異なっているが、共有設定に任意のタグを設定することで、地図表示時に同一タグを使用した質問の結果を組み合わせることができる。（この調査では、共有タグ名は `katasu` としている）。このため、調査回数を重ねるごとに、自動的に地点密度の高い地図が出力される。

使用に関する質問の選択肢は、カタスについては「a：言う」「b：聞いたことがある」「c：聞かない」の3段階とした。現時点では、同じタグがついていると、選択肢が異なる質問も統合されてしまう問題がある。統合するかどうかは表示時に選択可能だが、どの調査がどの選択肢で実施されたかは表示されないため、今後は選択肢に関する処理が必要になると思われる。

回答と地図記号の対応は、言語地図の出力時に決定することができる。授業でプロジェクターを用いる場合はカラー出力が効果的であるが、紙で配布する場合はモノクロで出力する必要があるため、異なる記号を割り当てることができる。現在は記号の種類は少ないが、本稿のカタスの選択肢は3段階であるため、「a：■」「b：○」「c：+」の記号を割り当てた。

地図記号は生育地にポイントされる。生育地の質問文は「5～15歳までに最も長く居住した場所」となっている。回答者の転居歴を考慮せずにすべて表示しているが、今後は転居歴によって表示するかどうかを選択できるようにする必要があるだろう。

7.3 調査結果

RMSシステムを用いたカタスの使用に関する調査結果を図7～9に示す。首都圏の大学での調査であるため、生育地は関東南部に偏っており、東京の周辺部の地図で示す。図7が2011年6月に初めてRMSの実験をおこなった授業（2クラス）での結果、図8が2012年に国立国語研究所プロジェクトで実施した統一調査での結果、図9が2011年6月から2013年6月までに調査したすべての結果を重ね合わせたものである。

図7をみると、カタスは首都圏のほぼ全域で使用されているように見える。不使用者（「o：聞いたことがある」「+：聞かない」の回答者）もいるものの、分布の傾向は不明である。この状態では首都圏ではカタスが再普及し、普及は完了したとってしまうだろう。

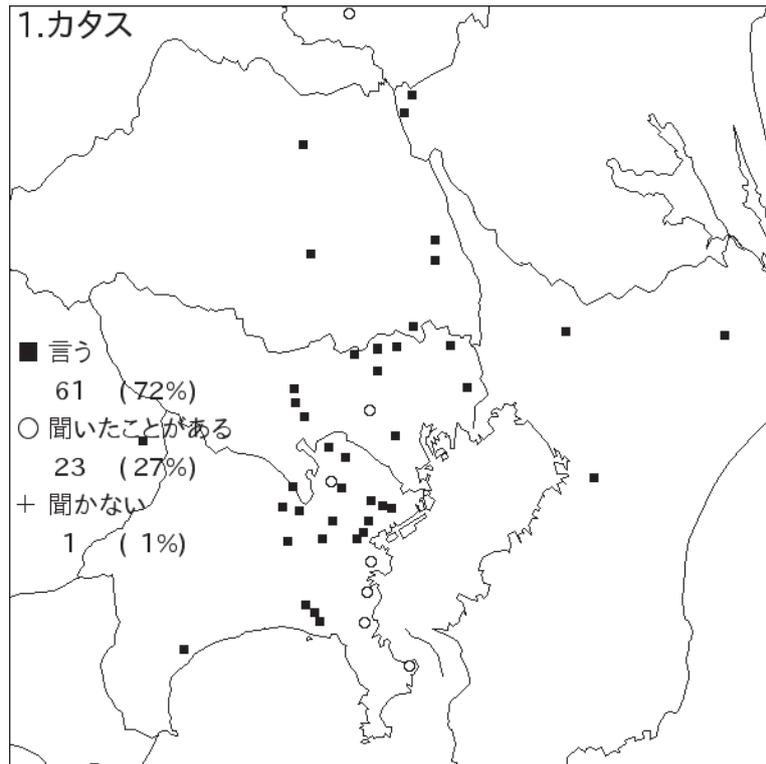


図7 カタスの分布 (2011年の調査)

つづいて図8をみる。2012年7～11月に国立国語研究所共同研究プロジェクトのメンバーで共同で実施した調査の結果である。調査では、携帯メール以外にも、質問紙を用いた調査も実施している。RMSシステムでは、携帯電話以外で実施した調査についても、結果を読み込むことで統合することができる。

地点数が増加したことにより、カタスは首都圏で完全には普及しておらず、神奈川県や、埼玉県の西部、東京都の西部に不使用者が多いことがわかってきた。先行研究において、神奈川県で使用されていなかったことや、千葉県でカタスが再普及中であったことをあわせると、カタスは関東東部から南西部に普及したと予想できる。しかし、図8では東京都多摩地域や神奈川県のデータが少ない。特に東京都の多摩東部地域に使用者がいないようにみえるが、本当にそうなのか判断しにくい。

図9はこれまでの調査結果すべてを重ね合わせたものである。2年間(3年度)で、9大学32授業の1834名の回答者を対象とした。全員が首都圏生育者ではないが、それでも首都圏内の地点はかなり高密度になっており、分布も解釈しやすい。特に東京都の多摩東部地域でカタスが不使用者であることが明確となり、カタスが普及する過程で、東京都多摩東部への普及が遅れていることがわかった。

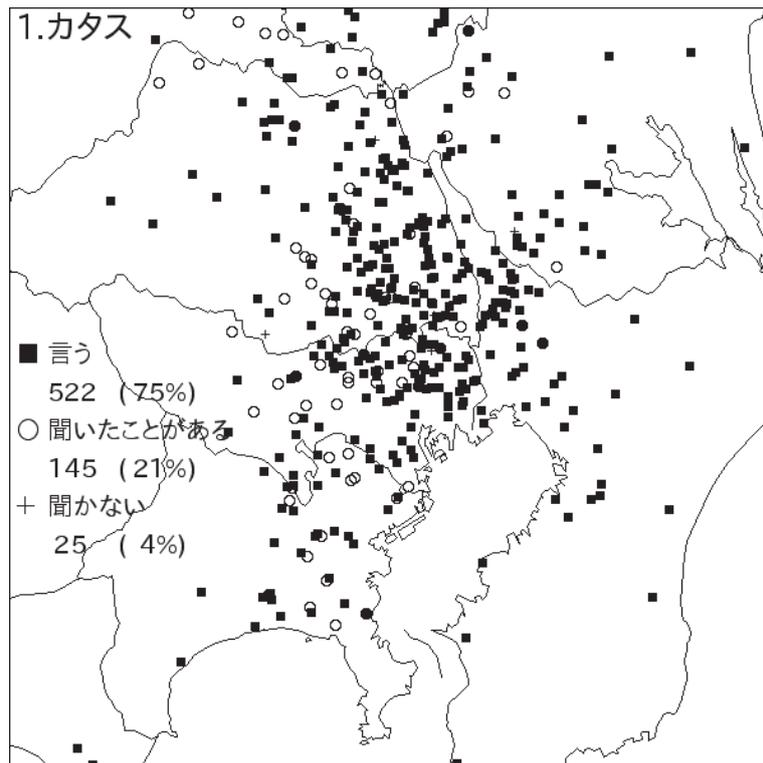


図8 カタスの分布 (2012年のプロジェクトでの調査)

なぜこのような分布になったかについて、1枚の共時的な地図だけでは、分布の形成過程がわからない。そのため、国立国語研究所の共同研究プロジェクトのRMS調査では、カタスの使用意識（使用頻度、通用範囲、通用場面、丁寧度）についても調査を実施した。その結果、鎌水・三井（2013）では、意識項目にも地域差が存在することがわかった。現在は、東京都の多摩東部地域だけが不使用地域になっているが、意識項目の地域差から、南に隣接する神奈川県や、東に隣接する東京都23区南東部も、カタスの普及が遅かったということが推定できるとした。

このように、首都圏において高密度の地点で表示することで、言語の動態を示すことが可能となる。1回の授業調査は40名程度であることが多かったが、RMSは授業中に短時間で実施でき、その場で結果を示すことが可能であるため、複数の大学、複数の授業において、効率よく効果的なデータを収集することができたといえる。

カタスについては2年間の調査データを重ね合わせて表示したが、短期間に使用状態が変化する場合、重ね合わせるのではなく、調査時期ごとに比較したほうがよい。RMS調査は短期間で広範囲に高密度のデータが得られるため、継続的に調査を実施することで、実時間の言語変化を地理的に観察することができる。RMSシステムは、変化の速い首都圏若年層の言語動態を把握するのに適したツールだといえることができる。

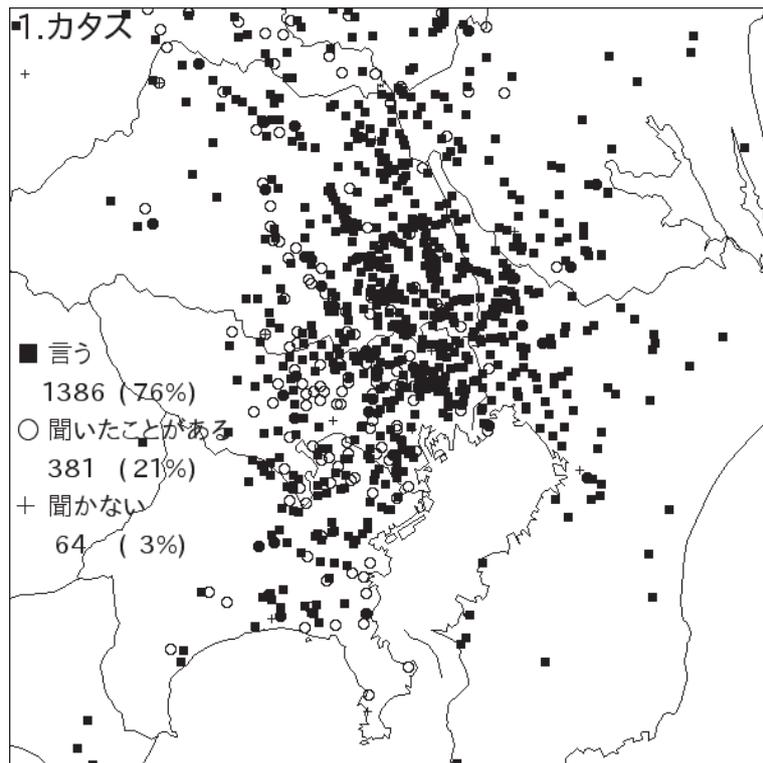


図9 カタスの分布 (2011年6月～2013年6月の全調査)

7. 4 授業におけるRMSシステムの利用

RMSシステムを利用することで、授業時間中に調査だけでなく、結果の提示や解説までをおこなうことができた。これ自体は意義があることだが、実際の授業においては、受講生数が少ないこともある。1回の授業で得られる回答が少ない場合は、結果が明確でなく、調査に対する理解も深まらないおそれがある。そのため教室内の調査結果を示したあとで、他の授業(他の大学)での結果を組み合わせた調査結果についても示すことが効果的と思われる。多くの場合、教室内的の結果が、周辺地域も含めた結果につながっていることを示すことができるため、学生の参加感を損なうことなく、調査の理解に役立つと思われる。

首都圏の言語は地域性が薄く、首都圏の内部は言語的に均質だと意識されていると述べたが、首都圏の大学において同じ教室内であっても言語的な均質性が完全でないことを認識させることは、首都圏における方言学の教育の面で効果的である。方言調査に対して興味をもたせ、また調査に協力しようという意識をもたせることにつながるとと思われる。

統計的に示すことはできないが、RMS調査に参加した学生の反応には、「新鮮だった」「すぐ結果がわかってよい」「楽しかった」といった肯定的評価が多かった。その一方で、教室内では少数派となる首都圏外の出身者からは「表示されるのは首都圏ばかり」「全国の結果が知りたい」といった不満もあった。結果を示す時間も限られており、授業における工夫にも限界がある。全国地図を常に画面上に表示するなど、システム面からの工夫も必要であろう。

また、言語地図の表示については、現在は教員が学生にプロジェクターで見せることを前提と

しているが、今後は学生側でも自由に地図を閲覧し、見やすいように加工できるようにすることも必要であると思われる。RMS システムでは、言語地図の記号を自由に選択できるため、学生が利用することで、言語地図の作成や分析の学習にもなる。このような教育ツールとしての改良も、今後の検討課題であろう。

8. おわりに

本稿では、首都圏若年層の調査において携帯メールを用いた調査を実施する意義と、開発したリアルタイム携帯調査システム (RMS システム) の概要と利用例について解説をおこなった。

インターネットの普及にともなって、調査・分析の方法も大きく変化している。携帯端末の普及は、各個人がデータ入力端末をもつことになるため、機動力のある調査に向いている。

面接調査は今後も重要であるが、従来の質問紙によるアンケートについては、今後インターネットを利用した調査が増加すると思われる。新しい、より言語実態を反映することができる調査手法の開発は、今後の方言研究の発展に寄与することであろう。

RMS システムは現時点では試験段階であり、国立国語研究所の共同研究プロジェクトのメンバー以外は、システムにアクセスできない。しかし、授業での利用という点で効果のある調査システムであり、研究者間でのデータの共有という点でも重要だと思われる。試験の協力者を拡大しつつ、最終的にはシステムの公開利用を予定している。

今後も RMS システムを改良・活用して、首都圏若年層における多様な言語状況の解明に役立てていきたいと考えている。

文献

- 福嶋秩子 (1983) 「パソコンによる言語地理学」『国語学』133, 105–106.
- 橋元良明 (2004) 「オンライン調査による社会調査の問題点」『日本語学』23(8), 明治書院, 180–190.
- 早野慎吾 (1996) 『首都圏の言語生態 (地域語の生態シリーズ関東篇)』, おうふう.
- 林良雄・日高水穂 (2008) 「Google Maps を用いたことばのアンケートシステム」『秋田大学教育文化学部研究紀要 (自然科学)』63, 21–25.
- 井上史雄 (1985) 『関東・東北方言の地理的・年齢的分布 (SF グロットグラム)』東京外国語大学言語学研究所.
- 井上史雄 (編) (1988) 『東京・神奈川言語地図』東京外国語大学井上史雄研究室.
- 井上史雄 (1994) 『方言学の新地平』明治書院.
- 井上史雄 (1997) 『社会方言学資料図集—全国中学校言語使用調査 (1993・1996)—』東京外国語大学.
- 井上史雄 (2000) 『日本語の値段』大修館書店.
- 金本良嗣・徳岡一幸 (2002) 「日本の都市圏設定基準」『応用地域学研究』7, 1–15.
- 河西秀早子 (1981) 「標準語形の全国的分布」『言語生活』354, 筑摩書房, 52–55.
- 木村治生 (2009) 「学校通しによる質問紙調査の可能性と限界」『社会と調査』2, 社会調査協会, 28–34.

- 岸江信介 (2007) 「パソコンで作る方言地図」 小林隆・小西いずみ・三井はるみ・井上文子・岸江信介・大西拓一郎・半沢康 『方言学の技法』, 岩波書店, 91-133.
- 国土地理院「地球地図日本」 <http://www1.gsi.go.jp/geowww/globalmap-gsi/download/index.html> (2013年7月31日参照)
- 国立国語研究所 (編) (1966-74) 『日本言語地図』 東京, 大蔵省印刷局.
- 前川喜久雄 (1988) 「PC9801 上の言語地図作成支援システム, EGL」 第 15 回近畿音声言語研究会.
- 内閣府 (2011) 「補論 1 首都圏人口の変化の長期的推移」 『地域の経済 2011』
http://www5.cao.go.jp/j-j/cr/cr11/pdf/chr11_4-1.pdf (2013年7月31日参照)
- 中井精一 (2005) 『社会言語学の調査と研究の技法』 東京, おうふう.
- 荻野綱男 (1975) 「コンピュータによる方言地図の作成と解析」 『第 20 回日本方言研究会発表原稿集』 30-39.
- 荻野綱男 (1981) 『方言調査分析用パッケージプログラム GLAPS の使い方』 私家版.
- 荻野綱男 (2004) 「電子メールによる質問調査法—調査の具体的な方法とその特徴—」 『日本語学』 23(8), 明治書院, 168-179.
- 岡本義雄 「日本列島海岸線データ& 県境データ」
http://www.cc.osaka-kyoiku.ac.jp/~yossi/programs_trash.html (最終閲覧日, 2013年7月31日)
- 大西拓一郎 (2002) 「言語地図作成の電算化—『方言文法全国地図』 第 5 集を例に—」 『日本語学』 21(11), 明治書院, 21-35.
- 大西拓一郎・鎌水兼貴・三井はるみ・吉田雅子 (2011) 『方言の形成過程解明のための全国方言調査—方言メール調査報告書—』 (国立国語研究所共同研究報告 10-02), 国立国語研究所.
- 柴田武 (1956) 「言語形成期というもの」 石黒修・泉井久之助・金田一春彦・柴田武 (編) 『子どもことば』 (ことばの講座 6), 東京創元社, 243-266.
- 総務省統計局 (2012a) 「平成 22 年国勢調査 大都市圏・都市圏 全国図」
http://www.stat.go.jp/data/chiri/map/c_koku/daitoshi/pdf/2010.pdf (2013年7月31日参照)
- 総務省統計局 (2012b) 「平成 22 年国勢調査 従業地・通学地による人口・産業等集計結果 結果の概要」 <http://www.stat.go.jp/data/kokusei/2010/kihon4/pdf/gaiyou.pdf> (2013年7月31日参照)
- 総務省統計局 (2012c) 「第 1 表 男女別人口及び世帯の種類 (2 区分) 別世帯数—全国, 市部, 郡部, 都道府県, 20 大都市」 <http://www.stat.go.jp/data/kokusei/2010/kihon1/pdf/gaiyou2.pdf> (2013年7月31日参照)
- 高橋頭志 (2003) 「『全国大規模言語地図』 調査・作成システム—スグダス言語調査システムと SUGDAS.exe—」 『国語学会 2003 年度春季大会予稿集』 203-210.
- 高橋頭志 (2007) 『日本語方言の層位—GAJ-Sugdass2006—』 群馬県立女子大学文学部国文学科高橋頭志研究室.
- 田中ゆかり (2004) 「質問して何がわかるか—言語調査法入門—」 『日本語学』 23(8), 明治書院, 6-18.
- 田中ゆかり (2006) 「『新規登場事象』をとらえたい 探索的調査法による多人数調査」 伝康晴・田中ゆかり (編) 『方法』 (講座社会言語科学 6), ひつじ書房, 55-72.

- 徳川宗賢・山本武（1967）「電子計算機で言語地図を作る試み」『計量国語学』40, 27-30.
東京大学空間情報科学研究センター「シンプルジオコーディング実験」
<http://newspat.csis.u-tokyo.ac.jp/geocode/modules/geocode/>（2013年7月31日参照）
- 鐘水兼貴（2007）『方言文法地図』における共通語化の状況—多変量解析を用いた分析—『日本語学』26(11), 明治書院, 112-119.
- 鐘水兼貴（2011a）「携帯電話を利用した首都圏若年層の言語調査」『情報処理学会研究報告〔人文科学とコンピュータ研究会〕』2011-CH-92, 1-13.
- 鐘水兼貴（2011b）「多様な方言資料を統一的に扱うための検索システムの開発」『日本語学会 2011年度秋季大会予稿集』263-270.
- 鐘水兼貴（2012）「携帯電話を利用したリアルタイム方言調査システム」『日本行動計量学会第40回大会抄録集』349-352.
- 鐘水兼貴・小西いずみ・松丸真大（2010）「方言調査データのXMLによるデータベース化」『情報処理学会研究報告〔人文科学とコンピュータ研究会〕』2010-CH-88, 71-81.
- 鐘水兼貴・三井はるみ（2013）「首都圏若年層における非標準形使用意識の地理的分布」『第31回社会言語科学学会研究大会発表論文集』162-165.