

## リサーチ・メソッドに関する一考察

山崎 友子

## はじめに

人と人との触れ合いの中で確かめられ温められていく教育という営みに携わるものにとって、「経験」は宝物である。しかしながら、経験を積みば積むほど、自らの経験に足をとられ、それを対象化することが難しくなることもある。そこで、実践を対象化する手段としてリサーチの有用性に着目し、2000年度後期大学院英語科教育学特別演習を「リサーチ実践の場」とした。

現役の中学校教員である畠山さんが受講生である。私の学部授業「ライティングⅡ」を対象として、リサーチ・クエスチョンを立て、データを収集・分析し、論文を書く、という一連の作業を協同で行った。同時進行中で生身の人間を相手とする実践をリサーチするには、さまざまな配慮が必要であり、また、その限界もしっかりと認識しておく必要がある。畠山さんとの協同リサーチをもとに、留意点をまとめてみた。

本稿は畠山さんとの協同論文「岩手大学生のパラグラフ・ライティング」の姉妹編としてお読みいただけると幸いである。

**仮説検証法と仮説生成法 (Hypothesis-testing vs. Hypothesis-generating Method)**

箕浦 (1999年) によると、どの実証研究も「データをどこから得てくるか」と「データのカテゴリゼーションをいつ、どのようにしてやるか」に証拠方式を設定しなければならない。仮説検証研究では、まず始めに仮説があり、その仮説にそったカテゴリーを頭において、データ収集のための研究がデザインされる。データ収集はカテゴリー化の後で行われる。得られたデータは、予め念頭にあったカテゴリーに従って量的に分析され、仮説が支持されるか否か検証される。研究は直線的に進められるので、ここから新しい理論が生まれることはないが、研究者として訓練を受け一定レベルに到達することが可能な研究方法である。SLA (第二言語習得) 研究には、このタイプの研究方法をとるものが多い。

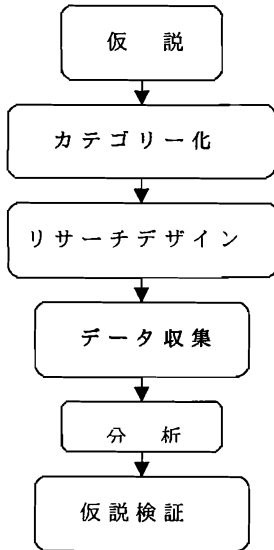


図 1 . 仮説検証法

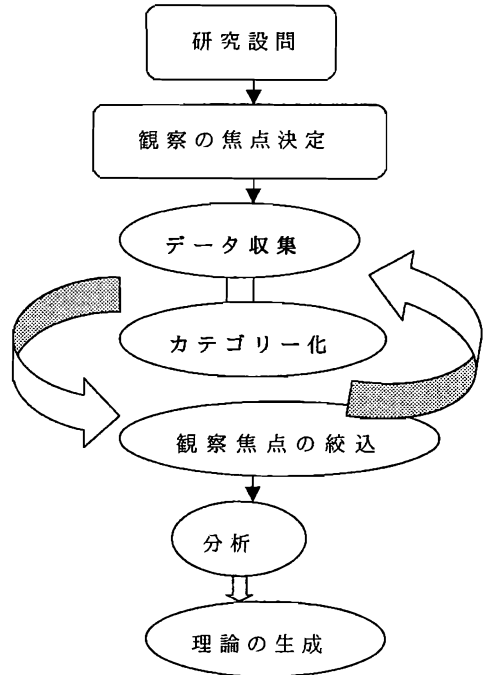


図 2 . 仮説生成法

仮説生成研究では、研究設問に対応して観察の焦点を定めデータが集められる。このデータ分析にあたって、現象を説明するのに最もふさわしいカテゴリーが先行文献などを参考に選択される。このカテゴリー化の作業は、データ分析と同時進行で行われ、観察の焦点がさらに絞り込まれる。そこで得られたデータは再度新たなカテゴリー化が必要かどうか検討される。これらの作業は長い時間をかけて、質的分析をもとにらせん状に進められる。観察する目の確かさ・カテゴリー化する判断の的確さが命である。研究者自身が研究のツールであるため、訓練によりすべてを習得できるというわけではない。しかし、この研究方法からは「データに根ざした」新しい理論が生まれる可能性がある。この方法論については B. G. Glaser & A. L. Strauss 『データ対話型理論の発見：調査からいかに理論を生み出すか』(1967年) に詳しい記述がある。

これら二つの研究方法を、後者は未開地を開墾するようなもの、前者は開墾された土地に種を蒔き、その収穫量を測定するような研究との箕浦（前出）のたとえは、両者の特徴をよく表している。

### 妥当性 (validity) と信頼性 (reliability)

Thomas K. Crowl (1993)によると、妥当性と信頼性は次のように説明されている。

**Validity:** The extent to which a test or measuring

instrument actually measures what it claims to be measuring. (テストや測定の道具が測定すべきものを測定しているかどうか、その程度)

**Reliability:** The extent to which a test or measuring

instrument measures something consistently.

(テストや測定の道具が一定した測定結果を示す程度)

たとえば言えば、体重を量るのに身長計を使っているか、象の体重を量るのに料理用計量器を使っているか、またその計器は電池がきれいかかっており、計るたびに違う数値を示すことはないだろうかということである。

協同研究「岩手大生のパラグラフ・ライティング」を例にとれば、English Proficiency を測定する道具として TOEFL と Illinois Vocabulary Inventory を用いた。これらはすでに長年にわたって数多くの受験者を対象に実施されている標準的なテストであり、採点のプロトコールが明確で信頼性は高い。しかし、妥当性に関してはどうであろうか。TOEFL から抜粋した文法問題では特にプレテストで Part B の得点率が極めて低かった。これは単に時間不足だけでなく、問題が被験者の英語力に照らして難しすぎたためかもしれない。だとすると、英語力を測定するものさしとして妥当性が低かった可能性がある。しかし一方で、奇数問題をプレテストに用い、偶数問題をポストテストに用いることによって難易度の一貫性 (internal consistency) を保つことができ、誰が採点しても変動のない信頼度の高いデータを得ることができるという利点がある。

このように妥当性と信頼性を高める作業には両面性があることが少なくない。複雑な現実を解く道具に完璧なものが存在することは少なく、仮説検証法においては検証すべき課題を明確に絞り、それに見合う道具を選択すること、仮説生成法においては道具である研究者自身を訓練や先行研究を学ぶことによって磨くことなどの工夫・努力が必要である。その上で、いずれの場合もリサーチの妥当性・信頼性を脅かすもの (threats) を最小にするデザインと、道具の持つ限界について意識的になることが求められる。

### 実験群と統制群 (Experiment / New Treatment Group vs. Control Group)

仮説検証法では、実験群と統制群が必要となる。たとえば、「オーラル・アプロ

一チが書く能力の伸長に有効である」との仮説を持っていて、その仮説を検証するのであれば、「オーラル・アプローチを実践する」だけでは不十分である。そこから生まれた成果が、本当にオーラル・アプローチから生まれたのか、あるいは他の要素から生まれているのかを確かめるために、全く同じ内容をオーラル・アプローチではなく他の教授方法を用いて教えたグループと比較する必要がある。そこで、「オーラル・アプローチによる授業実践」を行うグループを『実験群』、「オーラル・アプローチ以外の教授法による授業実践」を行うグループを『統制群』と呼んで実験を開始し、仮説を検証することとなる。

しかしながら、正規の授業を対象として『実験群』『統制群』を設定することは、実務的にも倫理的にも困難であることが多い。そこで、すべての学習者に同じ実践を行い、この実践の『前』と『後』でデータを収集し仮説を検証する方法がとられることもある。しかし、たとえ『実験群』と『統制群』を設定したとしても、仮説として立てている項目以外をすべて同じ条件にすることが可能であろうか。ガラスの試験管の中で条件を管理しての実験と異なり、教育実践の研究は意図せぬ事柄の影響から逃れることはできない。

できるだけ目的外の要素が入り込まないようなデータ収集の方法をデザインするとともに、結果の分析にあたって、他の要因の影響という可能性を慎重に考慮する必要がある。そこで、次の項では研究の信頼性を脅かす内的・外的要因 (internal/external threats) のいくつかを取り上げ、その対応の仕方を述べる。

### 量的分析・質的分析 (Quantitative vs. Qualitative Analysis) と阻害要因 (internal/external threats)

量的分析は仮説検証法に欠くことのできない分析方法であり、質的分析を中心とする仮説生成法においても必要に応じて用いられる。本稿では、主に量的分析の阻害要因を述べ、その補完手法としての質的分析の重要性に触れたい。

#### 1) サンプルング

量的分析に用いられるデータの妥当性・信頼性のチェックには統計的手法が用いられる。量的分析の妥当性を確保するには、何よりもサンプルングが重要である。研究対象の集団が母集団を代表するものとするには、もっとも一般的にはランダム・サンプルングが行われる。言語習得分野ではランダム・サンプルングは可能であろうか。アンケート調査などでは可能であるが、長期的に実験・観察を継続する場合は無理である。したがって、論文ではどのような集団を対象とした研究かをできるだけ詳細に記述することが望ましい。

ランダム・サンプルングができない場合は、実験そのものに妥当性を脅かす要

素が入り込む危険性 (internal threats) があることを承知しておかなければならない。長期にわたる研究の場合、ボランティアを募ることもある。その場合、被験者は母集団全体の標準より研究に理解がある可能性が高く、より効果を示す結果となることが考えられる。また、ボランティアに謝礼が支払われる場合は (金額にもよるが)、被験者集団に社会経済的なバイアスがかかる可能性を考慮しなければならない。

## 2) 有意な差

現実にあるクラスをそのまま実験対象とする場合、『実験群』の成績と『統制群』の成績を実験前に調査する。もし、この段階で差があれば、集団の性質を示す差なのかどうか統計的に確かめる必要がある。(t検定)有意な差が認められた場合は、二つの集団の実験終了時の調査 (ポストテスト) 結果を比較するのではなく、それぞれの実験開始時と終了時の差 (gain) を比較することによって妥当性を維持できる。しかしながら、ある教授法が成績上位者には有効であるが下位者には有効でない、あるいはその逆のこともありうるので、実験開始時に集団に差が認められた場合は、仮説の内容によっては、妥当性が危ぶまれる。

## 3) 道具としての「人」

仮に実験開始時の『実験群』と『統制群』に有意な差が認められなくても、教授者が異なっている場合、教授者の教え方・個性などが結果に影響を与えることが考えられる。これは外的阻害要因 (external threats) の一つである。防止策としては、実験期間の前期と後期で教授者を入れ替え、internal consistency をチェックする、事前に教授法のトレーニングを行い、どのクラスもガイドラインに沿って等質の授業ができるようにしておくなどが考えられる。

また第2言語データの rating にも外的阻害要因が入り込む恐れがある。分析対象の言語を第1言語とする rater と第2言語とする rater では、注目点が必ずしも同じとは限らず、これ自体が研究テーマとなる程興味深いものである。したがって、どちらの rater に依頼するのが研究の目的に合致しているかを検討し、複数の rater に採点を依頼して、その internal consistency をチェックして信頼性の高いデータを得ることが望ましい。協同研究「岩手大生のパラグラフ・ライティング」では、英語特有のパラグラフ概念を身に付けることを目標とした実践が研究対象であったので、英語を第1言語とする二名に rater となり、学生の書いた作文を採点していただいた。

## 4) 実験外の出来事 (historical events)

また、実験期間中に学校ではさまざまな出来事が起こる。「岩手大生のパラグラフ・ライティング」を例にとると、他の選択科目の中で「パラグラフ・ライティン

グ」が教えられたり、海外研修旅行が実施されたりすると、分析データに影響を与えることが考えられる。量的分析を中心とする研究であっても、アンケートやインタビュー等により被験者のバックグラウンドや実験にたいする反応をつかんでおくことは、外的阻害要因の影響を配慮する上で有益である。

#### 5) データ分析の統計処理にあたって

データの分析に統計的手法を用いることによって、研究の妥当性・信頼性を高めることができるが、その解釈によっては逆に現実から遠ざかってしまう恐れもある。注意すべき基本的な事柄を以下に述べる。

##### ① 分布

データを処理する際、まず平均点に目が行きがちであるが、統計による各種の検定は正規分布を想定して作られているので、データの分布をまず点検することが重要である。分布は平均値 (mean)・中央値 (median)・最頻値 (mode) の比較、標準偏差 (SD)、箱型図 (box plot)、葉状図 (stem and leaf) にまず目をやる習慣をつけるとよい。はずれ値は単に例外的な数値として切り捨てるのではなく、個別に検討するとこれだけでも興味深い情報をもたらしてくれる。

##### ② 相関関係 (correlation)

相関関係は統計では相関係数  $r$  (correlation coefficient) で表される。教育実践において相関関係を見る場合は、 $r$  の数値は多少低めでも相関関係があるとみなすことが現実的である。しかし、相関係数が高いとそこに因果関係 (cause and effect) があると誤って解釈しがちである点には十分気をつけなければならない。たとえば、「岩手大学生のパラグラフ・ライティング」の中で、Reading の伸びと Writing の伸びの間には比較的高い相関関係が見られた。 $(r = 0.35)$  この数値は「Reading が伸びている学生ほど Writing の伸びも大きい傾向が見られる」という事実を示しているのであって、「Reading が伸びれば Writing も伸びる」という因果関係として解釈してはならないのである。

##### ③被験者の減少 (attrition)

実験を行う上で最大の難関は被験者を確保することであるといっても過言ではない。ようやく集めた被験者も気まぐれであり、あらゆるデータをすべての被験者から集めることは困難である。特に、ポストテストで被験者が減少することがある。このような場合、すべての実験に参加した被験者のデータのみを使用することが必ずしも最善とは限らない。たまたま都合が悪くて受けられなかった場合もあるが、実験の効果が薄い被験者が脱落していくこともままあるので、脱落の理由をできるだけ把握してその処理を決定する必要がある。

「岩手大生のパラグラフ・ライティング」では、後期授業の開始週に出席し、

同意書を提出してプレテストを受験した学生は30名であったが、第二週以降には約50名の学生が受講することになった。畠山さんやアメリカ人コーダーからのフィードバックという被験者へのお返しは50名全員に与えられたが、同じ条件でプレテストを受験しなかった学生のデータは残念だが使わないこととした。

### 研究倫理

とりわけフィールド・ワークを主とする文化人類学・民俗学の分野で、得られたデータは誰のものか、ということが論議された。「調査地被害」という用語も生まれた(宮本 1972)。被験者によって体験され語られたことが、研究者によって論文という名の商品となっていく過程には、「倫理」という名の購入手続きが必要だとする問題提起であった。

これらの批判の中には、個々の研究者のマナーというレベルで解決がつくこともあるが、それ以外に「情報公開」「インフォームド・コンセント」といった学問体系の中に綱領として位置付けるべき性質のものもあり、無頓着であることは許されない。

研究倫理は二つの段階で問われる。研究を開始するときとその成果を発表するときである。仮説検証法による実験を実施する場合は、被験者からの同意を得ることは比較的容易である。しかし、それでも同意書をとることによって、ホーソーン効果(Hawthorn Effect)が生じることが考えられる。得られた結果は純粋に実験された項目からではなく、新しいことを体験しているという高場のためかもしれない。逆に、身構えられ、本音での対応が得られないかもしれない。そこで研究者は、データの信頼度を高めることを選択するか研究倫理を優先するかの岐路に立たされることになる。参与観察の場合は、より一層問題は深刻である。

次の岐路は成果の発表の場合である。もし、第一の岐路で同意を得ていない場合、ここで被験者から同意を得ることは倫理上不可欠な手続きとなる。しかし、データがすでに収集されたことを後で知らされることになる被験者の思いが心地よいものでないことは想像に難くない。

学校はデータの宝庫である。少し手を伸ばせばすぐに興味深いデータが手に入る。これらのデータを利用するにはどのような手続きを踏めばいいのだろうか。アメリカ合衆国では、州あるいは学区、あるいはそれぞれの研究機関により、本人(未成年の場合は保護者)から、時によっては学校長・教育委員会からも予め同意を得ることが定められている。

日本ではまだ明確な規定が出来上がっていないわけではない。しかし、信頼性についての危惧は他の手立てによって最大限補うこととし、研究開始の段階で総括

的な同意を得るようにするのがいいのではないだろうか。参考までに、資料1にアメリカ合衆国で中学生を被験者としてリサーチの際使用した同意書を載せた。個人名を出さない約束、何を行ってもらうのか明確に伝えること、結果は希望者についてでも知らせる準備があることを述べることをポイントとした。

日本ではこのような同意書を取る習慣がないため、署名をすることによって生じるのが研究者の義務であるにもかかわらず、被験者の義務ととられる心配もある。書面以外での補足説明や、口頭で同意を得ることなど、状況に応じて工夫の必要があろう。

### おわりに

教育実践研究には実験的研究 (intervention-type) と記述的研究 (descriptive-type) がある。前者においては、研究の妥当性・信頼性を脅かすさまざまな阻害要因が存在する。だから限界の多い有効性の小さな研究だと考えるのではなく、現実を見つめ、変える可能性を模索する果敢な挑戦として価値のあるものだと考えたい。

今回の「岩手大生のパラグラフ・ライティング」においても、小さな一つの数値を得るために、大きな努力と時間を必要とした。しかし、自らの実践から得られたデータは生きた声となって、研究者でありかつ実践者である私達の耳に届く。現役の教師である院生には、今回の研究とこの方法論に対する考察を参考に、ぜひそれぞれの実践の場で教育実践を科学する試みに挑戦していただきたいと思っている。

### 資料1

#### To the Volunteers and their Parents

I am a student at XXX (the name of the school). My interest is Teaching English as a Foreign Language in Japan. I am planning to do my research on the development of communication as part of Professor Y's course. My research will be a pilot test to explore how Japanese people communicate in English and how they do in English. I hope the result will be able to suggest something to improve the education of Japanese learners of English. I would appreciate it if you would participate in my project.

I will be administrating questionnaires and interviews in (date). I assure you that your names will be kept confidential. A summary of the result will be made available to any volunteer who is interested in it.



I hope you will indicate your agreement by completing the attached consent form. There is no direct benefit to you for participating in this study, other than a small (postcard or something you would give them) treat from me in appreciation. If you have any questions, please feel free to ask me.

Thank you very much for your cooperation in this matter. I truly appreciate it.

The name of the researcher  
How to contact the researcher

----- Detach here -----

### Consent Form

We, \_\_\_\_\_/ \_\_\_\_\_(names), agree to participate in the project designed by (the name of the researcher) as part of Y's course at XXX(the name of the school).

Parent's Signature: \_\_\_\_\_

Date: \_\_\_\_\_

### 参考文献

- グレーザー、シュトラウス (Glaser, B.G & A.L. Strauss) 後  
藤隆他訳 1967年『データ対話型理論の発見：調査からいかに理論を生み出すか』新曜社
- 佐伯胖・松原望一編 2000年『実践としての統計学』東京大学出版会
- 箕浦康子編著 1999年『フィールドワークの技法と実際』ミネルヴァ書房
- Crowl, T. K. 1993. *Fundamentals of Educational Research*. Iowa: Brown & Benchmark.
- Nunan, David. 1992. *Research Method in Language Learning*. New York: Cambridge University Press.

(岩手大学教育学部英語教育講座)