HOKKAIDO UNIVERSITY

Title	質問紙の回答の不安定性を引き起こす要因 ： 学習者ビリーフを調査する質問紙を使用して
Author(s)	小池, 真理
Citation	北海道大学留学生センター紀要, 6, 37-52
Issue Date	2002-12
Doc URL	https://hdl.handle.net/2115/45627
Type	departmental bulletin paper
File Information	BISC006_004.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

北海道大学留学生センター紀要　第6号　(2002)［研究論文］

質問紙の回答の不安定性を引き起こす要因[1)]

－学習者ビリーフを調査する質問紙を使用して－

小　池　真　理

要　旨

第二言語習得研究、言語教育などの量的研究においては多項選択式の質問紙を使用した調査が多く行われている。しかし、その質問紙の回答は質問項目の構成、質問文の言葉遣い、回答形式、回答者の認識など多様な要素により影響を受け、不安定になることがある。したがって、調査を行なう際には、質問紙を作成する段階で不安定性の要因を減少させることに注意を払うと同時に、その特質を把握して結果を分析することが重要であると考える。そこで、本研究では学習者ビリーフを調査する質問紙を取り上げ、その回答の不安定性の要因を分析し、質問紙調査時の注意点を提言することを目的とする。

30名の日本語学習者を対象に1～2週間の間隔をあけて同じ質問紙による2回の調査を行なった。調査の結果、質問紙への回答の不安定性には、①質問文の言葉遣い（ワーディング）の問題、②回答者が想定した状況、③回答者の意見、考え方の強さ、④回答傾向が複合的に作用していることが示唆された。これらの不安定性は完全に排除できるものではなく、調査結果の分析の際に配慮が必要となる。

〔キーワード〕質問紙調査、不安定性、学習者の回答傾向、ワーディング、学習者ビリーフ

1．はじめに

第2言語習得研究、また言語教育などの文脈においては、質問紙調査が多く実施されている。簡便性、汎用性に優れ、大量のデータにより一般化した議論が可能となるからである。しかし、その簡便さの裏には、どんな質問紙調査によっても何らかの回答が得られる、という不確実性も含んでいる。元来人間が考えていることを質問紙で調査することには限界があり、その限度を踏まえた上での調査である。したがって、その限界の中で信頼

性と妥当性を高く維持し、質問紙の特性を把握して測定したいことを調査する必要がある。質問紙には、質問項目の構成、質問文の言葉遣い（以下ワーディングとする）、回答形式、回答者の回答時の環境など多様な要素が含まれており、それらの要素が回答に影響を及ぼしている。事前に質問項目の構成、ワーディングに関して十分な検討が必要なことは言うまでもないが、それでもなお高い安定性があるとは断言できない。

学習者ビリーフ研究の文脈において使用する質問紙の安定性に関しては、Sakui & Gaies（1999）のように統計学的に検証を行なっている研究も僅かではあるが存在する。だが、安定性を検証して日本人英語学習者のビリーフを明らかにすることが目的であるため、回答の不安定さがどのような要因によって引き起こされるのかといった分析は行なわれていない。したがって、このような研究はほとんど行なわれていないと言っていい。本稿は、学習者ビリーフを調査する質問紙を使用して、その回答の不安定性を分析するものである。質問項目の構成は各調査者が目的に応じて考慮することであるので、ここでは論じない。

2. 目　的

5段階評定式質問紙調査における回答の不安定性要因を探ることを目的とする。学習者ビリーフを調査する質問紙を使用して、1～2週間の間隔をあけた2回の調査を行い[2]、その回答の安定性を分析する。そして、不安定であるものに関してその要因を分析し、質問紙調査の際の注意点を提言する。

3. 方　法

3.1 使用した質問紙

質問紙は小池（2002）によるものを使用した。この質問紙は、日本語学習者の視点及び社会調査法における質問文作成時の注意点[3]に留意してある。社会調査法の観点を取り込んだのは、ワーディングによる回答の不安定性を減少するためである。質問紙は以下の様に作成されている。まず、教育現場の日本語教師三人の内省により構築した質問項目にHorwitz（1987）が構築したBALLI（Beliefs About Language Learning Inventory）[4]の質問項目を訂正追加して、質問項目をリストアップした。次に、予備調査として七人の上級日本語学習者に対してリストアップした質問項目を使用

した調査、及び日本語の上達要因、学習者ビリーフ、回答形式に関するインタビューを実施した。そして予備調査の結果から、質問項目の再構成を行い、指摘された理解しにくい語彙、表現などの改良を加えた。また、漢字には全て振り仮名を付け加えた。

質問紙は100の質問項目からなり、その回答形式は「1.強く賛成する」、「2.賛成する」、「3.どちらとも言えない」、「4.反対する」、「5.強く反対する」の5段階評定である。

3.2 データ収集

調査の対象者は、国立大学の留学生センターで日本語コースを履修している大学院生と研究生である。日本語の質問紙が理解できる日本語レベルの学習者を選定した。日本語レベルの判定基準は、中級の中以上の日本語クラス[5)]を履修していること、また大学院入試で日本語論述の試験を受けて入学したことなどを総合的に判断した。

学習者53名に調査の目的などを説明後、質問紙を直接手渡した。1～2週間後に回答を回収し、再度質問項目の提示順を変えた同質問紙を手渡した。さらに1～2週間後に回答の回収を行なった。1回目の回収は38名分であり、2回目は25名分であった。これとは別に留学生以外で質問紙を理解できる学習者4名と上記クラスの修了者3名に質問紙を郵送した。1～2週間後に1回目の回答が返送されてから2回目の質問紙を郵送した。2回目の回答は5名分回収した。以上合計で1回目は44回答、2回目は30回答であった。

最終的なデータは30名分であり、対象者は男性が13名、女性が17名である。学習者の母語は英語、ロシア語、中国語、韓国語など14言語で、全員日本語以外の外国語の学習経験がある。

4. 各質問項目の安定性の分析

ここでの安定性の分析は、1回目と2回目の測定値の相関によって行なうのではなく、個々の回答者に関して2回の回答が一致したかどうか、また同じ意見であったかどうかによって行なうものである。

4.1 安定性の高い質問項目と低い質問項目

2回の回答が一致した回答者数の割合を算出し、一致回答者率（CAR）

とした。さらに、同傾向の回答を選択した回答者数、つまり「強く賛成する」と「賛成する」のように強さに関係なく同じ意見を選択した回答者数の割合を同傾向回答者率（SAR）とした。（CAR と SAR は筆者が新たに採用した指数である。）

$$\text{一致回答者率（CAR）} = \frac{\text{2回の回答が一致した回答者数}}{\text{全回答者数}}$$

$$\text{同傾向回答者率（SAR）} = \frac{\text{2回の回答で同様の意見を選択した回答者数}}{\text{全回答者数}}$$

以上のように一致回答者率と同傾向回答者率を算出し、値の高い質問項目と低い質問項目を表4に記した。つまり、表4は安定性の高い項目と低い項目を表している。最高値は、一致回答者率が82.1％で、同傾向回答者率が96.4％であった。また、最低値は、一致回答者率が25.0％で、同傾向回答者率が48.1％であった。表4は、一致回答者率が75％以上の項目と同傾向回答者率95％以上の項目の上位群と前者が40％以下と後者が50％以下の下位群をまとめた。

表4　安定性の高い項目と低い項目

	質問項目	M1*	M2*	CAR %	SAR %
安定性の高い項目	46. 日本語を勉強するためには、日本文化を知る必要がある。	2.00	1.96	82.1	89.3
	48. 教師から文法規則の明確な説明を受けることが重要である。	2.07	1.93	82.1	92.9
	51. 教室の外で日本語を話すとき、分からない語彙や表現があったら、聞き返した方がいい。	1.79	1.86	78.6	96.4
	30. 日本語だけでなく、態度や表情などもコミュニケーションには重要である。	1.79	1.64	78.6	92.9
	21. たくさん読めば、読み方が上達する。	1.75	1.79	78.6	92.9
	72. 正しい発音で話すことが重要である。	1.82	2.00	78.6	92.9
	17. 漢字を覚える時、熟語で覚えた方がいい。	2.39	2.22	78.6	78.6
	7. 実際の日本語の使い方を学習するべきである。	1.64	1.86	67.9	96.4
	12. 日本語母語話者と同じように話せるようになりたい。	1.71	1.71	71.4	96.4

	25. 自分の日本語の誤りがわかるようになることが重要である。	1.46	1.71	60.7	96.4
	20. 書いた文の誤りを先生に直してもらって学習することが重要である。	1.46	1.54	64.3	96.4
安定性の低い項目	6. 学習した文型は書いて覚えることが重要である。	2.36	2.11	25.0	53.6
	4. 特に学校に通っていなくても、日本語を聞いたり話したりしていれば自然に覚える。	2.82	2.86	28.6	53.6
	53. 聞き分けられない音は発音できない。	2.61	2.93	35.7	53.6
	81. 読む時は、一つ一つの語彙を理解して、文をつなげて読むことが重要である。	2.78	2.93	37.0	48.1
	2. 日本語を始めた最初の時期に誤りが訂正されないと、その誤りはずっと残って後で直すのが難しくなる。	2.30	2.32	37.0	63.0
	73. 聞き分けられても、発音できない音がある。	2.79	3.14	39.3	50.0
	38. 漢字を覚える時、例文で覚える。	2.46	2.79	39.3	50.0
	22. 他の言語が話せる人は日本語学習が簡単である	3.57	3.25	39.3	53.6
	10. 一つ一つの音を正しく発音することが最も重要である。	2.39	2.75	39.3	57.1

*M1は1回目の調査の平均値、M2は2回目の平均値を示す。

この表から分かることは、以下の5点である。ただし、①と②に関しては本稿では議論しない。

①安定性の高い質問項目には、(48)と(20)の教師の役割、(12)の動機付け、(51)、(30)、(7)、(72)の実際のコミュニケーションに関わる項目、(25)の自己モニターがある。また、スキル別の学習ストラテジーの中でも(17)の漢字、(21)の読解が安定している。さらに、(46)のような言語学習における文化の認識に関わる項目が安定している。

②安定性の低い質問項目には、(6)、(73)、(10)のような音声に関わる項目、(22)のような一般的な人に関して問う項目がある。

③安定性の高い質問項目は、評定値の平均がほとんど1点台で、賛成傾向が強い項目である。

④安定性の低い質問項目は、評定値の平均が3点前後で、「賛成、反対のどちらとも言えない」という意見に近い項目である。

⑤安定性の低い質問項目には、(4)、(2)、(22)のような学習全般に関する項目よりその他の項目のようなスキルに関わる質問項目が多い。

安定性が低い質問項目に関して、上級学習者へのインタビューで得られたコメント、及び質問紙の回答に付記されたコメントを分析した。[38.漢字を覚える時、例文で覚える] に関して「具体的にどのような方法を言っているのかわからない」という質問があり、この学習ストラテジーに関して認識がなかった可能性がある。[6.学習した文型は書いて覚えることが重要である] に関して「自分のノートに書くという意味ですか」という質問があり、実際の活動が想定しにくい表現であることが示された。このことから、回答者によって異なった状況を想定した可能性があると言える。

また、③と④の結果から安定性の高い質問項目に対しては明確な強い意見を持っているのに反して、安定性の低い質問項目に対しては明確な強い意見を持っていず「どちらとも言えない」との間で意見が揺れていることが推察できる。予備調査の結果から、回答者は迷わずに賛成か反対を選択する場合と、しばらく考え迷いながら選択する場合とがあることがわかった。日常意識的に持っている考え方とあまり意識していない考え方では、回答時の迷い方にも差が生じるであろう。

以上の分析から、不安定性の要因として意見の強さと学習者が想定した状況が挙げられる。

4.2 意見の強さの差が大きい質問項目

表4の安定性のある項目の(25)や(20)、及び安定性のない項目の(6)や(4)のように一致回答者率と同傾向回答者率の差が大きい質問項目がある。このことは、2回の回答で、同意見の異なった強さを選択した回答者が多いことを示している。そこで、一致回答者率と同傾向回答者率の差が30.0%以上の質問項目を表5に挙げた。これらのビリーフの強さは、流動的であると言える。

表5 意見の強さの格差が大きい質問項目

質問項目	M1	M2	差 %
8. 婉曲的な表現を学習しなければならない*。	1.79	1.89	39.3

25. 自分の日本語の誤りが分かるようになることが重要である。	1.46	1.71	35.7
27. 学習した文型は会話の中で覚えることが重要である。	1.75	1.93	35.7
71. 学習した文法、表現、語彙または、テレビや町の中で聞いた文法や表現、語彙はすぐ使ってみたほうがいい。	1.82	1.89	35.7
78. 日本語は、書くことより読むことのほうがやさしい。	2.04	2.11	33.3
20. 書いた文の誤りを先生に直してもらって学習することが重要である。	1.46	1.54	32.1
40. 文章を書く前に、全体の構成を考えてから書いたほうがいい。	1.79	2.14	32.1

*実際の質問紙では(8)に会話例が提示してある。
(例：A：明日、映画に行きませんか。B：すみません、明日はちょっと…)

意見の強さの格差が大きい質問項目には、ビリーフの強さが関わっていると言える。これは、予備調査での質問紙回答時に、「絶対、そう思う」と即座に「強く賛成する」を選択した場合と迷いながら「やっぱり強く賛成するかな」と選択した場合があることから分析できる。例えば［71.学習した文法、表現、語彙または、テレビや町の中で聞いた文法や表現、語彙はすぐ使ってみたほうがいい］では、「使ってみることは重要だけど、状況によりますね」というコメントがあり、重要性は認めてもどんな状況でも重要であるわけではないと考えていることが示唆された。また、［40.文章を書く前に、全体の構成を考えてから書いたほうがいい］では、「そう思うんだけど、実際にはあまりやっていないから。どうしたらいいかな」というコメントがあり、回答に迷っている様子が覗えた。一方、一致回答者率が高い［46.日本語を勉強するためには、日本文化を知る必要がある］や［48.教師から文法規則の明確な説明を受けることが重要である］(表４)では、ほとんど迷うことなく「強くそう思う」や「文化はそんなに必要じゃない」などを選択していた。以上のことから、ビリーフの強さは絶対的なものと相対的なものがあると考えられる。絶対的なものは個人が常に意識的に保持しているビリーフであり、相対的なものは、他のビリーフとの相対的な関係で強さが決まるビリーフである。ただし、絶対的なものであっても従来意識されることが少なかったビリーフでは強さが流動的になる可能性がある。したがって、意識的に保持している絶対的なビリーフは２回の回答が一致する可能性が極めて高いはずである。相対的なものと意識

化されることのないビリーフは、一致率が低下し5段階評定での強さが流動的で回答時によって強さに変化が生じやすいと考えられる。

また、これらの質問文に使用した文型は初級で学習した文型であること、また語彙は「婉曲的」「構成」以外何度も質問紙内で使用されているものであることから、語彙や文章の意味がわからずに不安定になったということは考えにくい。「婉曲的」に関しては会話例を示してあること、さらに対比して「直接的な表現」を使用した質問文で同様に会話例を示してあることから、語彙による影響は少ないと考える。ただし40.の「構成」に関しては語彙の影響である可能性は否定できない。

ビリーフの強さは予備調査のインタビューや質問紙の回答に付記されたコメントから分析したものであるが、ここで挙げた質問項目の不安定さの要因を意見の強さの一つに限定したわけではない。

4.3 変化率の高い質問項目

4.1で、2回の回答が一致したか否か、また同じ意見か否かを分析して安定性を分析した。しかし、5段階評定における評定値の変化は、1と2、4と5のように同意見の場合と、2と3、3と4のようにどちらでもないを含んだ場合、さらに2から4のように全く異なった意見を選択した場合とがある。そこで、明らかに異なった意見が選択された質問項目を分析する。個々の質問項目に対して、1回目と2回目の測定値が2以上の差を示した回答者数の割合（変化回答者率UAR）を算出し、上位群を表6に挙げた。（UARは筆者が新たに採用した指数である。）

$$\text{変化回答者率（UAR）} = \frac{\text{2回の回答の測定値の差が2以上の回答者数}}{\text{全回答者数}}$$

表6　変化率の高い質問項目

質問項目	M1	SD1	M2	SD2	UAR %
94. 教室の外で日本語話すとき、直接的な表現をあまり使わないほうがいい*。	2.50	1.00	2.50	1.04	35.7

6. 学習した文型は書いて覚えることが重要である。	2.36	1.03	2.11	0.69	21.4
81. 読む時は、一つ一つの語彙を理解して、文をつなげて読むことが重要である。	2.78	1.09	2.93	0.98	18.5
26. 日本語学習でいちばん重要なことは、語彙をたくさん覚えることである。	2.86	0.88	2.68	0.92	17.9
49. たくさん文法規則を勉強すれば、話すことも上達する。	2.21	1.11	2.57	0.98	17.9
53. 聞き分けられない音は発音できない。	2.61	1.10	2.93	1.12	17.9
73. 聞き分けられても、発音できない音がある。	2.79	1.20	3.14	1.01	17.9
99. 発音できても聞き分けられない音がある。	2.82	1.22	2.82	1.12	17.9
59. たくさん読めば、書くのが上手になる。	2.71	1.15	2.64	0.95	17.9

*実際の質問紙では94に会話例が提示してある。
(例：A：明日、映画に行きませんか。B：すみません、行きません。)

ここで94に注目したい。変化回答者率は35.7％であり、約三分の一の回答者で異なった意見が選択されている。しかし、これは表4の安定性の低い項目のリストには、挙げられていない。すなわち、「一致」を基準とした場合、特別に安定性が低いわけではないにも拘わらず、「変化」を基準とした場合には変化回答者率が高い。さらに、変化回答者率の高さに反して、評定平均値は一致している。これは、不賛成から賛成に変わった回答が6、賛成から不賛成に変わった回答が4あったことに起因する。したがって、この質問項目は異なった意見を選択した学習者が多いにも拘わらず、従来の再テスト法による信頼性の検証や評定平均値を使用したt検定や分散分析では不安定性が見えないということである。

4.4 t検定の結果との比較

表4の安定性の低い項目、及び表5と表6の項目に関して、平均値の差を比較するt検定を行なった。その結果、5％水準で有意差が認められたのは［25.自分の日本語の誤りがわかるようになることが重要である］($t=2.260$) であった。しかし、これは表5の意見の強さの格差が大きい質問項目に含まれるものの、表6の変化率の高い項目や表4の安定性の低い

項目に含まれないばかりか、表4の安定性の高い項目に含まれる質問項目である。つまり、統計的検定で有意差が生じたものの、これは同様の意見であるが異なった強さを選択した学習者が多かったことを示しているにすぎず、不安定であるとは言えない。

従来の研究（細田1994、岡崎1996、板井2000）では、質問紙の安定性の検証をせずに評定平均値を使用したt検定、分散分析によってビリーフ変容や差異を論じている。しかし、上記のように平均値だけでは測定できないものが存在することから、従来の研究方法には不備があると言える。例えば、[94.教室の外で日本語話すとき、直接的な表現をあまり使わないほうがいい] のように平均値に差がない限り、質問項目に対して各回答が不安定であっても、統計的検定ではなんら影響を受けない。逆に [25.自分の日本語の誤りがわかるようになることが重要である] のように安定性のある項目でも、意見の強さに大きな変化があるため評定平均値に差が生じて、t-検定で有意差が認められることもあり得る。統計的検定で有意差が認められなくても、そもそも各学習者においてそのビリーフが不安定なものであれば、安定性があるとは言えない。評定平均値を使用した信頼性検証だけでなく、詳細な安定性の分析及び個々のデータ分析が必須である。

5. 回答者による安定性の違いの分析

Bachman & O'Malley (1984) は、質問紙のリカートスケールにおいて、アフリカ系アメリカ人(黒人)のほうがアングロサクソン系アメリカ人(白人) より、5段階評定における1と5の回答を多く選択する傾向にあると述べている。また、Hui & Triandes (1989) は、同様にヒスパニックとヒスパニックではない人を比較して、前者のほうが後者より1と5が多い回答傾向を示すことを報告している。上級学習者へのインタビューで5段階評定の回答形式に関して質問した結果、1と5を選択するのが難しいと述べた学習者は4人、問題なく5段階で回答したと述べた学習者は3人であった。7人の回答の傾向には個人差が見られた。以上のように背景文化の違いや個人間により回答傾向に差があると予測できる。そこで、回答傾向の違いにより回答の安定性に差が生じるかを分析した。

5.1 回答者の回答傾向の違いによる安定性の違い

5段階評定のうち「強く賛成」・「強く不賛成」の選択頻度が高い回答者

（以下、高頻度Ｌ）また、選択頻度が低い回答者（以下、低頻度Ｌ）に関して、回答の安定性を比較した。表７のように、高頻度Ｌでは、２回の回答の一致率が53.4％であった。低頻度Lでは一致率が64.1％で、高頻度Ｌより10ポイント以上一致率が高くなっている。また、５段階評定で測定値に２以上の差がある項目の割合、すなわち１回目と２回目で異なった意見を示した項目の割合である変化率に関しては、高頻度Ｌが9.0％で、低頻度Ｌが4.86％と一致率と反対の傾向が見られた。ここで、低頻度Ｌの標準偏差が高い値を示しているが、これは全体のばらつきが大きいのではなく、１名の回答者の変化率が高いことによる。この値を削除すると、表中の括弧内の数値が示すように低頻度Lの変化率は3.3％で、標準偏差は2.6となる。

表７　「強く～」の選択頻度と一致率

	「強く」選択数	一致率	SD	同傾向率	SD	変化率	SD
低頻度Ｌ	12.8	64.1%	10.2	76.0%	9.0	4.9 (3.3)	6.4 (2.6)
高頻度Ｌ	39.8	53.4%	5.4	77.7%	5.7	9.0%	4.5
全　　体	25.4	59.1%	9.8	76.8%	7.6	6.8%	5.9

さらに、低頻度Ｌと高頻度Ｌを独立変数、一致率、変化率を従属変数としてカイ二乗検定による有意差検定を行なった。「強く賛成」「強く不賛成」の選択頻度、一致率、変化率をそれぞれの平均の近似値でリコードした。その結果、一致率は$\chi^2=5.14$（Ｐ＜.05）であった。また変化率は$\chi^2=7.34$（Ｐ＜.05）であった。したがって、一致率、変化率とも５％水準で有意差が認められた。「強く賛成」「強く不賛成」の選択頻度が高いときは、一致率が低く、変化率が高い傾向がある。しかし、同傾向率はほぼ同じ値を示し、意見の強さが流動的であると言える。これは、表５の意見の強さの格差が大きい質問項目に影響を与えていると考える。

５段階評定において１や５の選択頻度が高いと、一致率が低く変化率が高い傾向にあることが明らかになった。しかし、同傾向率は、選択頻度が

低い回答者と変わらない。したがって、同傾向率が高く、一致率が低い質問紙は、質問紙自身の安定性の問題というより、学習者の回答傾向、ビリーフの強さ、想定した状況、などの要因の複合的作用によると考えられる。

6. まとめ

以上の調査結果から質問紙の回答の不安定性を引き起こす要因として、①質問文のワーディング、②回答者が想定した状況、③回答者の意見、考え方の強さ、④回答者の回答傾向があり、これらが複合的に作用していると考えられる。質問文のワーディングに関しては質問紙作成時に十分な配慮が必要であるが、注意を払って作成した質問文においてさえ完全な安定性を確保することは不可能である。

また、回答者の回答傾向の影響もあるため、特に異文化間の人に対する質問紙調査で比較研究する際には、単純に評定平均値を比較して意見や考え方の強さの違いを議論することはできないと言える。

以上から質問紙調査時の注意点を以下に挙げる。

①質問紙作成時にワーディングに注意を払う。
注3）にも記載したが、質問の意図を明確にし、曖昧な表現を避け簡潔な表現を使用することが重要である。

②新しい質問紙を使用した調査をする際には、予備調査を行い回答者がどのように質問文を捉えているか把握しておく必要がある。

③②の予備調査を実施しない時は、Sakui & Gaies（1999）が提案しているように質問紙調査後のフォローアップインタビューを実施し、回答者の捉え方を調査する必要がある。

④汎用化した質問紙を作成する時、また縦断的研究により考え方の変容を調査する時は、予備調査での信頼性・妥当性の検証は不可欠である。さらに信頼性の検証の際に個々の質問項目の安定性も分析するべきである。特に縦断的研究においては予備調査を実施しなければ、本来不安定な考え方だったのか、強さが異なりやすい考え方だけか、またその期間の何らかの介入によって変容したのか見極められなくなる。

⑤調査結果を分析する際に、②、③から得た質問紙の特徴を把握し分析することが求められる。

本研究によって得られた知見は、5段階評定の回答形式の是非が問われる問題でもあるかもしれない。Hui & Triandes（1989）は10段階評定にす

ることにより回答傾向の違いの影響を減少できると報告している。しかし、回答者にとって10段階はより煩雑になり回答しにくさが増す可能性がある。また、3段階にするという考え方もあるであろうが、明らかに強い意見に関しての弁別力がなくなるともいえる。この回答形式の問題は今後の研究の課題である。

注：

1）本稿は北海道大学国際広報メディア研究科の2001年度修士論文（小池、2002）の一部である。

2）社会調査法での質問紙の信頼性の検証方法である再テスト法では、通常2～4週間の間隔をあけて調査が行なわれるが、本稿では、質問項目が多く記憶に残りにくいことを考慮し、かつ日本語クラスにおける教師のビリーフの影響を軽減するために、1～2週間の間隔とした。

3）社会調査法における質問紙作成時の注意点は安田（1987）、井上等（1995）、大谷等（1999）を参考にして関連する項目をまとめると、以下の通りである。

①質問の意図を明確にすること
- 回答の時点や内容を明確に指示する。
- 一般的質問なのか個人的質問なのか明確にする。

②文章を簡潔な表現にすること
- 1つの質問に2つ以上の論点を入れてはいけない。
- 否定法による文章による混乱を避ける。

③わかりやすい語句を用いること
- あいまいで多義的な言葉を使用しない。
- 難しい言葉を使用しない。
- ステレオタイプの言葉を使用しない。

④誘導的な言葉や文章を避けること
- 誘導質問にならないように気をつける。
- イエステンデンシー（黙従傾向：「～がいいと思いますか」と論点を一方向に提示すること）にならないように気をつける。
- 言葉や文章が回答者の反感や共感をあおらないように気をつける。

4）BALLIは34の質問項目からなり、「言語の困難さ」、「言語学習の適性」、「言語学習の特質」、「学習ストラテジーとコミュニケーションストラテ

ジー」、「動機付けと期待」の5カテゴリーに分類されている。回答は1の「強く賛成する」から5の「強く反対する」までの5段階評定式が32項目、a）からe）までの順序尺度が2項目である。

5）調査対象学習者は、中級の中以上の文法、作文、読解クラスを履修している。ただし、併行して会話クラス、漢字クラスを履修している学習者もいる。

参考文献

Bachman, J. G. & O'Malley, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black-white differences in response style. *Public Opinion Quarterly*, 48, 491-509.

Horwitz, E. K. (1987). Surveying student beliefs about language learning. Wenden, A., & Rubin, J. (edts.) *Learner Strategies in Language Learning*. London, UK: Prentice-Hall International, 119-129.

Hui, C. H. & Triandes, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, Vol. 20. No. 3, 296-309.

Sakui, K. & Gaies, S. J. (1999). Investing Japanese learners' beliefs about language learning. *System*, 27, 473-492.

井上文夫・井上和子・小野能文・西垣悦代（1995）『よりよい社会調査をめざして』　創元社 pp.80-116

板井美佐（2000）「中国人学習者の日本語学習に対するBELIEFSについて—香港4大学のアンケート調査から—」『日本語教育』104号、日本語教育学会 pp.69-78

大谷信介・木下栄二・後藤範章・小松洋・永野武（1999）『社会調査へのアプローチ』　ミネルヴォ書房 pp.66-91

岡崎　眸（1996）「教授法の授業が受講生の持つ言語学習についての確信に及ぼす効果」『日本語教育』89号、日本語教育学会 pp.25-38

小池真理（2002）「日本語教育における学習者ビリーフの調査法」北海道大学国際広報メディア研究科2001年度修士論文

細田和雄・伊藤克弘・本田直美（1994）「日本語学習者と日本語教師養成過程大学生の日本語学習に関するビリーフ」『広島大学日本語教育学科紀要』4号、広島大学日本語教育学科 pp.85-90

安田三郎・原純輔（1982）『社会調査ハンドブック』第3版、有斐閣双書
pp.136-141

Journal of the International Student Center Hokkaido University No. 6 (2002)

Factors Which Induce Instability of Questionnaire Responses
Using questionnaires to investigate learners' beliefs

Questionnaires using the Likert scale are often used in quantitative research in the context of second language acquisition and foreign language teaching. However, questionnaire responses are often unstable due to the influence of various factors, such as the construction of items, the wording of sentences, the scale style and how respondents understand the sentences. Therefore, in research using questionnaires it is very important to pay close attention to the wording of sentences, and also to analyse the data with the effect of these factors in mind. The aims of this study are to identify factors which induce instability of questionnaire responses and to make some suggestions for this kind of research.

Questionnaires were responded to by 30 participants twice, 1～2 weeks apart (the first time using the original version, the second time using a scrambled version). Findings of the analysis were that relevant factors were (1) the wording of question sentences, (2) respondents' assumptions concerning situations, (3) the strength of respondents' opinions and (4) a tendency to extreme responses, and that these factors affect instability in combination, not individually.