

# 日本語学習者の自己評価 Learner's Self-Assessment of Japanese Proficiency

小山 悟  
南山大学大学院

## 要 旨

本論文は、学習者に自己の学習プロセスを意識化させるための具体的な手段のひとつとして「自己評価」の導入を提唱し、学習者の自己評価の信頼性とその信頼性を高める要因について検討するものである。自己評価の信頼性については、学習者の自己評価と教師評価との相関を算出し、これを信頼性の根拠とした。また、自己評価の信頼性を高める要因については、まず「学習者の自己評価はモニター(m)活動によって形成される」との仮説を立て、モニター(m)に影響を与えると予想されるいくつかの要因の中から「日本での滞在歴」の影響を、その他の要因の中から「学習者の受けた学校教育のレベル」と「日本語の熟達度」の影響について分析を行った。調査の結果、自己評価と教師評価との相関はこれまでの先行研究の結果と比べて必ずしも有望なものではなかったが、統計的には0.5%水準から1%水準で有意であった。また、自己評価の信頼性に影響を与えると予想された3つの要因の中では「日本での滞在歴」でわずかにその影響が確認されたものの、その他の2つの要因では特に影響は見られなかった。その原因については「メタ認知ストラテジー」の観点から考察を加えた。

## 1 はじめに (Introduction)

第二言語習得や学習ストラテジーに関する研究が進んだ結果、学習者が言語学習に対してどの程度主体的に関わっているかが、言語学習を成功に導くうえで、重要な鍵を握っていることが明らかになってきた。田中・斉藤(1993)らによって、日本語教育における「自律的学習」の必要性が強く主張されるようになってきたのもそのためであろう。とはいえ、学習者すべてがはじめから自律的な学習者であるわけではないし、また自律的でない学習者に対して「自分の学習に責任をもて」と言うだけで、その学習者が自律的になるわけでもない<sup>(1)</sup>。学習者が言語学習に主体的に取り組むことができるようになるためには、少なくとも、学習者がメタ認知を効果的に働かせ、自己の言語学習プロセスを意識化できるようになる必要がある。そこで本研究では、学習者が自己の言語学習プロセスを意識化できるようにするための具体的な手段のひとつとして「自己評価」の導入を提唱し、学習者の自己評価の信頼性とその信頼性を高める要因について検討する。

## 2 研究の目的 (Purpose)

本研究の目的は、次の2点に要約される：1) 学習者の自己評価はどの程度の信頼性がある

か、2) 自己評価の信頼性を高める要因は何か。

第一の点については、従来の自己評価研究が一貫してテーマとしてきたことである。古くは Balke-Aurell (1977)<sup>(2)</sup> に始まり、その後 Oskarsson (1978, 1981) や LeBlanc and Painchaud (1981, 1985)、von Elek (1982) らによって、調査が繰り返して行われてきた。しかし、いまだはっきりとした結論には至っていない。その原因のひとつは、Backman and Palmer (1989) が指摘しているように、目的もデザインも異なる自己評価研究がバラバラに行われてきたことにある<sup>(3)</sup>。そこで本調査では、Oskarsson (1978) の開発した自己評価フォームを使い、Oskarsson (1981) と同じ調査方法で自己評価の信頼性について検証し、Oskarsson の調査結果との比較を行う。

第二の点については、これまでほとんど扱われてこなかったテーマである。これまで、自己評価の信頼性を高める要因として、自己評価フォームの形式や質問の内容については論じられてきたが<sup>(4)</sup>、そもそも自己評価をするということはどのようなプロセスを経るものなのかについては、全くと言っていいほど論じられてこなかった。自己評価研究がまだ結論に至らないもうひとつの原因がそこにあると思われる。そこで、Krashen のモニター理論をめぐる様々な議論を参考にして、自己評価に影響を与えると思われる要因について仮説を立て、それらについて検証することにする。

### 3 仮説 (Hypothesis) — 自己評価に影響を与える要因 —

#### 3-1 二つのモニター

自己評価とは何かについて考えるとき、Krashen のモニター理論をめぐる様々な議論は非常に重要なヒントを与えてくれる (Krashen, 1982, Krashen et al, 1984, Krashen and Terell, 1983)。

Krashen の理論の特徴は「学習」と「習得」を明確に区分したこと、そして意識的な学習の役割を極めて限定的に捉えたことにある。Krashen は意識的な学習によって得た知識は、習得された知識によって生成された発話に対するモニターもしくはエディターとしての機能しか果たさないと述べている。これは言い換えれば、モニターの機能を学習された知識による形式的な規則の運用としてのみ捉えているということを意味している。また、モニターを使用するためには文法規則を知っているだけでは不十分で、それ以外に 1) 十分な時間があること、2) 意味ではなく言語形式に焦点を当てていること、という二つの条件が揃っていなければならないと主張する。しかも、規則を知っているからといってその規則をモニターとして使えるとは限らず、モニターが行使されるのは低次の文法規則だけであるとも述べている。Krashen がこのように極めて限られた条件の中でのみモニターが行使されると考えたのは、モニター理論を裏付ける証拠を文法テストに見られる不自然な習得順序に求めたからである。

しかし、このような Krashen の限定的な解釈については、Morrison and Low (1983) ほか数多くの批判がある。Morrison and Low は、モニターとは Krashen の言うような言語形式に焦点を当てた明確に意識化されたものばかりでなく、直感によって感じる批判能力とも呼ぶべきものを含んでいると述べている。この点に関しては、Amato (1993) も「我々は常に潜在意識的にはモニターしており、習得した仮説と耳にしたり発話したことが一致しないことがあると、その時だけモニターしたと感ずるのではないか」と述べている<sup>(5)</sup>。Krashen も習得による自己訂正の機能について言及してはいるが、それがどのようになされるかについては

何も述べていない。また、Morrison and Low は Krashen のモニター理論では言語の産出は説明できても、言語の受容は説明できないと主張する。そしてまた、Krashenのモニター理論ではモニターを聞き手と話し手との協同作業とは捉えておらず、そのためモニター使用は言語の統語面に限定されているが、実際には発音や語彙、談話をも調整する能力をもっているとも述べている。これは言い換えれば、モニターとは、吉田(1991)が述べているように、Canale and Swain(1980)や Savignon(1983)らが提唱する、いわゆるコミュニケーション能力の構成要因である言語能力、ディスコース能力、社会言語学的能力およびストラテジー能力の各能力を総動員してなされるものだということであろう<sup>6)</sup>。Ellis(1988)は Krashen の限定的な意味でのモニターを「大文字(M)のモニター」と呼び、より広い意味でのモニターを「小文字(m)のモニター」と呼んでいるが<sup>7)</sup>、これに従えばKrashenのモニター(M)はモニター(m)に包含されることになる。

この二つのモニターについて、吉田(1991)は言語活動の違いによるモニターの仕方の違いについて指摘している。吉田によれば、産出力の場合には Krashen の言うようなモニター(M)が用いられ、言語の形式そのものに注意を払っている可能性が高いが、理解力の場合にはむしろより広く一般的なモニター(m)を利用して言語の形式以外の知識を活用している可能性のほうが高いという。また、同じスキルでも相手と直接話す場合には、聞き返しのストラテジーを使うなどしてある程度相手の言語表現をコントロールすることができる場合があり、その場合にはモニター(M)を使うことも可能であろうが、テープやラジオなどのように話し手が目の前にいない場合にはモニター(M)を使うことはほとんどできなくなる、というようにモニターの仕方に差が生じるとも述べている。

さて、自己評価とモニターの関係についてであるが、これまでの自己評価に関する先行研究の調査結果は、自己評価がより広い意味でのモニター(m)活動の結果であることを示唆している。例えば Oskarsson(1978)は、「学習者は人工的な学習環境における言語運用やテストのスコアの観点からではなく、自然な環境における言語使用の観点から、いつも自分の言語能力を評価しようとしている」<sup>8)</sup>と述べており、Wenden(1986)による学習者の自己意識調査もこれを裏付けている。Wendenは、上級レベルの第二言語学習者たちにインタビューを行い、学習者たちが実に様々な方法で自己の第二言語能力を自己診断していることを明らかにしている。ある学習者は以下のように述べている。

My English was much better than it was in England. I understood more and I talked more easily. I had made progress, but I couldn't measure it. It was all unconscious. But I was always trying to get information in a conscious way. I read the ads, listened to radio commercials and watched TV. (Wenden, 1986 : P.186)

また、LeBlanc and Painchaud(1985)や Peirce et al(1993)は、第二言語を使う具体的な状況を設定した質問が最も正確な自己評価結果を生むと報告しているが、これは学習者たちが日常の言語運用場面で無意識のうちにモニター(m)を使って自己評価を繰り返していることを裏付けているといえる。

Oskarsson(1978)は日常の言語運用場面で無意識のうちに繰り返される自己評価を「インフォーマルな自己評価」と呼び、教室内で学習活動の一環として行われる自己評価を「フォーマルな自己評価」と呼んでいるが、「フォーマルな自己評価」とはモニター(m)を使って無意

無意識になされる「インフォーマルな自己評価」を自己評価フォームという紙面の上に意識化するということを意味していると考えられる。

### 3-2 モニター使用に影響を与える要因

フォーマルな自己評価が日常の言語運用場面で繰り返されるインフォーマルな自己評価を意識化したものであるなら、モニター(m)の使用を促進する要因によって自己評価も影響を受ける可能性が大きい。そこで以下では、Krashen et al (1984)の指摘しているモニター(M)使用に影響を与える4つの要因(1.学習者の年齢、2.学習者の受けた言語教育のレベル、3.言語表現テストで求められる言語活動の性質と焦点、4.学習者の性格)がモニター(m)の使用にも影響を与えるかどうか検討する。

#### 3-2-1 学習者の年齢

思春期が第二言語の習得にとって一つの転換期になっていることがこれまでの研究によって明らかになっている。例えば、Asher and Garcia (1969)による第二言語の習熟度に関する調査では、当該言語が話されている土地における滞在期間の長さよりも、その土地に到着したときの年齢のほうが発音の習得に決定的な影響を与えることがわかっている<sup>(9)</sup>。調査によると、6歳以前にアメリカに到着した子供は最終的に母語話者と変わらないレベルまで到達することができるが、13歳以降に到着した子供は長期に渡って滞在している場合でも母語話者に近づく可能性は極めて低く、7歳から12歳までに到着した子供の場合には母語話者に近づく可能性は五分五分だったということである。

この思春期という年齢は、発達心理学でいう「形式的操作期(formal operational stage)」に当たる。ピアジェ(Piaget)によると、子供は形式的操作期(11歳以降)に達すると抽象的な概念や命題についても操作できるようになり、また仮説を立て、推測し、演繹的に思考することが可能になるとされている<sup>(10)</sup>。実際、第二言語習得の研究でも、形態素や統語に関しては少なくとも初期の段階では、成人のほうが年少者よりも速く習得するということがわかっており、この説を裏付けている<sup>(11)</sup>。

Krashen et al (1984)は、「言語について抽象的に考える能力、つまり言語理論や文法理論を構築したり、理解したりする能力は、形式的操作による思考能力に伴って発達する能力に、左右される」<sup>(12)</sup>と述べ、モニターが形式的操作期をひとつの転機として使われるようになることを指摘している。しかしこれはKrashenの言う意識的な「学習」、つまりモニター(M)に関することであって、これらの仮説がモニター(m)にも適用されるかについては明らかではない。ただ、多くの自己評価研究が大学生や成人学習者を対象にしている中、中学2年生を対象にして行われたPeirce et al (1993)の調査で、自己評価と客観テストとの相関が他の調査に比べて非常に低い数値を示したことから考えて、モニター(m)にも年齢の何らかの影響があると思われる。

#### 3-2-2 学習者が受けた言語教育のレベル

Krashen et al (1984)は学校での意識的な文法教育がモニター(M)使用に与える影響について、いくつかの調査結果を報告している。

Krashen et al (1976)は、教室でかなりの期間(3~15年)英語の訓練を受けたものの、教室外で英語との接触がほとんどない学習者と、学校でほとんど英語の訓練を受けていないもの

の、アメリカに3ヵ月から50年間住んでいる学習者を対象に、文法項目の習得順序に関する調査を行った。その結果、両グループとも極めて類似した習得順序を示したとのことであった<sup>(13)</sup>。Krashenらによれば、この結果は意識的に学習した文法規則がモニター(M)として活用されていないことを示すものであり、「意識的に規則を知るとは、モニター使用にとって必要なことであるが、このことよってのみ、学習者が表出することばに、モニターが適用されるとは限らない」<sup>(14)</sup>と述べている。また、Krashen et al (1978) は大学のESLの学生を対象に、作文における文法の自己修正に関する調査を行っている。調査の結果は、分析された6個の形態素に関する限り、自己修正できたのは全体の6%に過ぎなかったとのことであった<sup>(15)</sup>。このことは、学習者に自分の作文を慎重にチェックするように指示しても、極くわずかな量しか修正できないことを示している。これらの結果から、モニター使用のためには文法規則を知っていることが前提条件の一つで、その意味では学習者の受けた言語教育はモニター(M)使用に影響を与えはするものの、その影響の度合いについては必ずしも大きなものではなく、その範囲も限られているということになる。実際、文法や発音といったより多くのモニター(M)を活用すると思われるスキルについては、学習者はあまり正確な自己評価ができないことがこれまでの調査でわかっている<sup>(16)</sup>。

この傾向は、モニター(m)使用の場合にはどうであろうか。というのも、モニター(m)は意識的に学習された知識によるものではなく、Krashenも認めているように、無意識な「感覚」(feel)とも言うべきものによって行使されるからである。また、モニター(m)が前述の吉田の言葉の通り、コミュニケーション能力に類似したものであるなら、それは教室での意識的な文法学習によって得られるものではなく、むしろ日常の自然な場面での言語運用を通じて獲得されるものであると考えられるからである。とすれば、Dornic and Ekehammar (1988)が「言語能力の知覚は様々な個人的経験の統合体であることは明らかである」<sup>(17)</sup>と述べているように、モニター(m)は、学校での意識的な文法教育によってではなく、むしろ学習者の受けるインプットの量や、学習者の経験するインターアクションの量によって影響を受けると考えるべきではないだろうか。Krashen (1977) はモニター(M)について、インプットにさらされることによって開発されると述べているが、このことはむしろモニター(m)にこそ当てはまるのではないだろうか。

### 3-2-3 言語活動の性質と焦点

Krashen et al (1984) は、モニター(M)の使用に関して、「言語項目の操作に焦点を当てるテストは、モニター使用を促進し、意志伝達に焦点を当てるテストは、モニター使用を促進しないようである」<sup>(18)</sup>と述べ、「言語形式に焦点を当てる」ことがモニター(M)使用のための条件の一つであることを指摘している。では、モニター(m)の使用はどんな言語活動によって促進されるのであろうか。

自己評価フォームに関する研究から、学習者がより信頼性の高い自己評価ができるようにするための具体的方法の手掛かりとして、二つの興味深いことがわかっている。ひとつは、既に述べたように、学習者が自分の実体験と照らし合せて質問に答えることができるように、質問の内容や状況設定をできるかぎり学習者が現在置かれている状況に近付ければ、自己評価の信頼性を高められるという点である。もうひとつは、学習者は自分が簡単だと思うことよりも難しいと思うことの方が気づきやすいのかもしれないという点である<sup>(19)</sup>。この点については、先に紹介した「我々は潜在意識的には常にモニターしており、習得した仮説と耳に

したり発話したことが一致しないことがあると、その時だけモニターしたと感ずるのではないか」とする Amato の言葉を支持しているように思われる。また Schluë (1976) は、中級レベルの英語学習者について発話の自己修正に関する調査を行い、学習者がどのくらい文法的な誤りを修正できたか調べているが、結果はわずか10%程度の修正率であったという。そして、この結果について、自由会話のような伝達の緊急度の高い状況においては「学習者は文法的正確さよりも、自分が伝える内容の正確さと、談話のなかで自分が発する文の適切さに注意を向ける」と結論付けている<sup>(20)</sup>。

これらの結果から結論付けられることは、モニター(m) 使用は Krashen が言うような文法形式に焦点を当てた言語活動によって促進されるのではなく、日常の自然なコミュニケーション活動を通じて、あるいは意志伝達に焦点を当て、コミュニケーション能力を活用させるような言語活動によって促進されるということである。

#### 3-2-4 学習者の性格

Krashen et al (1984) は学習者の性格とモニター(M) の関係について、モニター使用者を 1) モニター過剰使用者、2) モニター過少使用者、3) モニター適用者の3つのタイプに分類している。それによると、モニター過剰使用者は「場独立型」<sup>(21)</sup> の傾向が強く、分析的な知覚スタイルを持ち、自意識の高い、内向的な性格の人が多いとされている。その反対に、モニター過少使用者は「場依存型」の傾向が強く、比較的自信があり、自分の誤りを気にしない、社会的かつ外向的な性格の人が多いとされている。

モニター(m) に学習者の性格がどのような影響を与えるかについては、これまで特に研究されてこなかったが、先の吉田(1991) が指摘しているように、モニター(m) がいわゆる「コミュニケーション能力」と密接なつながりがあり、意志伝達に焦点を当てた、コミュニケーション能力を活用させるような言語活動によって促進されるとすれば、「場依存型」の学習者の方がモニター(m) をより効果的に活用できるのではないかと考えられる。「場依存型」の学習者は Krashen の言うような「モニター過少使用者」などではなく、むしろモニター(m) の適用者である可能性すらあると断言していいのではないだろうか。

#### 3-3 自己評価に影響を与えるその他の要因

自己評価に影響を与えるその他の要因としては「学習者の第二言語の熟達レベル」や「学習者の受けた学校教育のレベル」などが考えられる。

##### 3-3-1 第二言語の熟達レベル

一般に学習が進み、その言語に関する知識が増えれば、それだけ正確な自己評価が可能になると思われるが、実際にはどうであろうか。これまでの自己評価研究の調査結果から、この推論を裏付ける根拠となるものを探してみることにする。

Wenden (1986) は前述の調査の中である学習者の以下のようなコメントを紹介している。

Now I am more conscious of what I do wrong than at the beginning. Now I know more English. I know that my problem is stress and sound. Before, I knew something was wrong [with her pronunciation] but I didn't know what it was.

(Wenden, 1986 : P.194)

このコメントは、学習の初期の段階で自己の問題点についてあまりはっきり認識できなかった学習者が、学習が進むにつれて徐々に具体的に問題点を認識できるようになったことを非常によく表している。

また、木山(1977)は日本の大学生や専門学校生を対象に英語の聴解力について自己評価の調査を行っているが、1回、2回、3回と調査を繰り返すうちに学習者たちの自己評価と正答との相関が高まったと報告している。

さらに、Yule et al (1987)も、聴解力の自己モニターに関する調査を行い、学習の進展にもなってモニター(m)の使用が改善されることを指摘している。それによると、聴解力は学習の初期の段階で急激に伸び、その後ある程度の段階にまで達すると、いわゆる「停滞期」に入るが、モニター(m)の正確さはむしろ停滞期に入ったころから徐々に改善され始めることが報告されている。

このように学習が進めば、より正確なモニター(m)、即ちより信頼性のある自己評価が可能になると思われる。しかしその一方で、後で詳しく紹介するが、Janssen-van Dienen (1989)の調査のように自己評価の信頼性と学習者の第二言語の熟達レベルとの間に有意な関係を見い出せなかったとする調査報告もあり、いまだ見解の一致は見られていない。この点については、今回の調査でも分析してみたい。

### 3-3-2 学習者の受けた学校教育のレベル

最後にこれまでのモニター(m)と自己評価をめぐる論議の中で特にその影響を指摘されてはいないが、自己評価の信頼性に何らかの影響を与えるのではないかとと思われる要因として、「学習者の受けた学校教育のレベル」を挙げたい。

Wangsootorn (1981)は、タイの大学生と大学院生を対象に自己評価の調査を行い、学習者の受けた学校教育のレベルの違いが自己評価に何らかの影響を与えているかどうか調べている。その結果、作文能力では両者の間に5%水準で有意差があることが判明したが、他の3つのスキルについては有意な差は見られなかったと報告している。しかし、信頼性のある自己評価ができるようになるためには、自己の言語学習活動をより客観的に把握することが必要不可欠で、そのためにはより高度な教育を受けた経験を持っていた方がよいのではないだろうか。調査項目の中に組み入れて分析してみることにする。

ここまでに検討した内容を改めて整理しておく、まず学習者の自己評価は日常の言語接触場面で無意識のうちに繰り返されるモニター(m)活動によって形成されると考える。そのモニター(m)は、ある程度の認知レベルに達し、より多くのインプットやインターアクションを経験した、場依存型の学習者によって、意志伝達に焦点を当てた言語活動においてより多く活用されると思われる。そこで、今回の調査では、この中から学習者の受けたインプットやインターアクションの量、即ち「学習者の日本での滞在歴」が自己評価の信頼性に影響を与えるかどうかについて検証を行うことにする。また、モニター(m)の観点からではないが、「学習者の受けた学校教育のレベル」や「学習者の日本語の熟達度」も何らかの影響を与えるのではないかと考えられるので、これについても分析を行うことにする。

## 4 調査の方法 (Research Methods)

調査は以下の手順で行われた。まず学習者ひとりひとりにレベル判定表を手渡し、自分の日本語力に関して4技能それぞれについて自己評価してもらった。同様に、各担当教師にもそれぞれが担当している学習者の日本語力を学習者と同じレベル判定表を使って評価してもらった。その後で、学習者に対して客観テストを実施し、学習者の自己評価を教師の評価と客観テストの成績のそれぞれと比較し、分析を行なった。分析の方法としては、相関やtテストなどの基礎的な統計手法を用いた。学習者による自己評価は客観テストの実施前に行ない、また学習者の自己評価と教師の評価もお互いの評価の結果に影響を与えないように別々に行なった。

### 4-1 被験者

調査の対象となったのは、名古屋市内の日本語学校に在籍し、大学もしくは専門学校への進学を主な目的として、日本語を学習している就学生82名である。そのうち、アンケート記入上不備のあったもの、明らかに質問の内容を取り違えたと思われるもの、出席状況が悪く教師が学習者の学力評価に困ったものを除いた66名が分析の対象となった。66名の内訳は、中国語話者56名(うち中国国籍52名、台湾国籍4名)、韓国語話者9名、ミャンマー語話者1名である。被験者の年齢は19歳から38歳までであるが、その大半(65名中53名)は20代で、平均年齢は25.5歳である。また、男女比では男性が約6割(66名中40名)を占めている。被験者の母国での最終学歴は、高校卒が39名(59.1%)と最も多く、以下大学卒15名(22.7%)、短大・専門学校卒12名(18.3%)の順となっている。被験者の日本での滞在歴については、大まかに二つのグループに分けられる。即ち、調査の実施時点でその滞在歴が6ヵ月から9ヵ月のグループと14ヵ月から16ヵ月のグループである。これは、現在の就学生を対象とした日本語学校が入学時期を4月と10月に設定しており、今回の調査が主に93年10月期に入学した学習者のクラス(初級終了レベル)と94年4月期に入学した学習者のクラス(中級レベル)を対象として行われたことによる。

### 4-2 自己評価フォームと客観基準

自己評価の質問表はOskarsson(1978)が考案したタイプC(添付資料参照)を使った。タイプCは英語で書かれていたものをまず日本語訳し、それを現在名古屋市近郊の大学、短大に在籍している台湾人と韓国人の留学生に翻訳してもらった。調査では中国語話者には中国語版が、韓国語話者には韓国語版が、ミャンマー語話者には英語版が手渡された。

比較のための客観基準には教師評価と客観テストが採用された。教師評価は、学習者と同じ質問表を使って、各クラスの担当教師に学習者の日本語のレベルを判定してもらった。客観テストは、学習者たちが当面の目標としている日本語能力試験1級の過去の問題から、聴解と読解の問題を使った。ただし、各設問の配点は公表されていないので、採点は筆者の任意の配点で行った。客観テストについては、実施段階でミスがあり、66名中31名しかデータを回収できなかった。また会話と作文については、利用可能な標準化されたテストがなかったので実施しなかった。



### 4-3 分析の方法

調査した結果は、tテスト、相関、一致率、一致度という4つの方法で分析された。

tテストは、二つのグループの平均値に差があるかどうかを検定する際、使われる統計計算法で、今回の調査では二つのグループの一致度や一致率に有意な差があるかどうかを検証するために使われた。

相関は、教師評価もしくは客観テストと自己評価との結びつきの強さを測るもので、これまでの自己評価研究で広く使われてきた手法である。今回の調査でも学習者の自己評価の信頼性を検証するために用いられた。

一致度と一致率は、自己評価と教師評価について、教師と学習者の評価のずれを調べるために用いられた。一致率は、全66ペアのうち学習者の自己評価と教師の評価とが一致したペアが何パーセントを占めているかを表している。例えば、66ペア中33ペアが一致すれば、一致率は50%ということになる。また、一致度については、教師と学習者の評価結果の誤差を以下に示すように絶対値で計算する。例のように一致度が0.53であれば、学習者の自己評価は教師の評価に比べて、平均0.53ずれていることになり、数値が小さいほど教師評価に近い自己評価を下したということになる。

	-2	-1.5	-1	-0.5	±0	+0.5	+1.0	+1.5	+2.0
人数	1	2	3	8	26	14	6	4	2

(例) 自己評価と教師評価のずれ

$$\{(2.0 \times 3) + (1.5 \times 6) + (1.0 \times 9) + (0.5 \times 22)\} \div 66 = 0.5303 \dots \quad \text{一致度 } 0.53$$

## 5 調査の結果 (Results)

### 5-1 自己評価の信頼性

まず、学習者の自己評価がどの程度正確で、どの程度信頼性のあるものなのかについて検証する。また、スキルによって自己評価の信頼性に差が生じるか、言い換えれば学習者にとって自己評価しやすいスキルはあるか、についても合わせて検証する。

従来の自己評価研究では、学習者の自己評価と他の客観評価(主に客観テストと教師評価)との相関を算出し、その結果を学習者の自己評価の信頼性を論じる上での根拠としてきた。しかし、そのような議論は客観的な比較基準である教師の評価や客観テストが学習者の能力を正確に判断しているという前提の上に成り立つものである。そこでまず、教師の評価と客観テストとの相関係数を算出し、教師の評価について客観基準としての信頼性を調べた。結果は、聴解については  $r = 0.784$ 、読解については  $r = 0.717$  であった。教師はかなり信頼度の高い評価を下していると言ってもよさそうである。会話と作文に関しては利用可能な客観テストがなく、その信頼性を検証する術がないので、聴解と読解の結果から推定する以外にないが、Blue(1988)が述べているように<sup>(22)</sup> 教師にとって難しいのは、学習者が口頭もしくは表記によって表現したものを評価する産出力よりも、むしろ学習者の内面を探って評価しなければならない理解力のほうであると考えられることから、会話と作文に関しても聴解、読解と同程度の信頼性は確保されていると考えたい。以下、教師評価を学習者の自己評価の信頼性を測る客観基準として採用することにする。

まず始めに、学習者の自己評価と教師評価との相関を調べてみた。結果は表1に示したとおりである。

分析の結果、学習者の自己評価と教師の評価との間には0.5%水準から1%水準で有意な相関があることがわかった。しかし、 $r=0.424$  という数値はこれまでの先行研究の結果と比較しても、必ずしも満足のいくものとは言えない。例えば、Oskarsson (1981) は本調査と同じタイプCを使って成人の第二言語学習者を対象に調査を行っているが、その調査で算出された相関係数は $r=0.77$ であった。また、Wangsotom (1981) はタイプCとよく似た自己評価フォームとしてWilkinsのレベル判定基準を使って調査しているが、その調査でも相関係数は最も低い数値を示したスキルでさえ $r=0.645$ であった。今回の調査に参加した学習者に関して言えば、その自己評価の信頼性は決して高くないと言えそうである。

表1 自己評価と教師評価・客観テストとの相関

	教師評価	客観テスト
会 話	0.284**	—
聴 解	0.381***	0.075
作 文	0.364***	—
読 解	0.309**	0.360*
総 合	0.424***	—

教師評価 df=64 \*\*\*0.5%水準で有意

客観テスト df=29 \*\*1%水準で有意

\*2.5%水準で有意

表2 自己評価と教師評価の一致度と一致率

	一致率 (%)	一致度 (P)
会 話	15.2	0.72
聴 解	34.8	0.61
作 文	22.7	0.77
読 解	22.7	0.72

次に、個々のスキルについても見てみることにする。既に述べたように、文法や発音に関する自己評価は学習者にとって非常に難しいという点では、先行研究はほぼ一致した見解を示している。しかし、その他のいわゆる4技能に関しては、それほどはっきりした見解は示されていない。例えば、Wangsotom (1981) や LeBlanc and Painchaud (1981) らが「産出力より理解力の方が正確な自己評価ができるようだ」と述べていたり、Blanche and Merino (1989) が「学習者は純粋なコミュニケーション能力のほうが自己評価しやすいようだ」と述べている程度である。また、Anderson (1982) や Dorric and Ekehammar (1988) のように、4つのスキルの自己評価は対等ではなく、学習者は発話能力や作文能力などの観点から自己の総体的な第二言語能力を評価しているのではないか、という意見もある。

今回の調査結果を見てみると、最も相関の高かったのは聴解(0.381)で、以下作文(0.364)、読解(0.309)、会話(0.284)と続いている。この結果から見るかぎり、産出力に比べて理解力の方が特に自己評価しやすいとは思われない。また、自己評価の正確さに関して、スキルによる違いや特徴らしきものがあるようにも思われない。

ところで、各スキルによる違いを相関とは別に一致率と一致度を使って分析してみた。結果は表2に示すとおりであった。

4技能の中で最も教師の見解と近かったスキルは聴解で、第2位は読解と作文であった。聴解は一致率が34.8%と最も高く、一致度が0.61ポイントと最も低かった。この結果を見ると、学習者には産出力よりも理解力の方がより正確に自己評価ができるように見えるが、算出された数値をtテストを用いて5%水準(両側検定)で検定した結果は、スキルによる自己評価のしやすさを結論付けられるほど有意なものではなかった。相関係数を使った分析の場合と同様、自己評価のしやすさに関してスキルによる違いや特徴があるようには思われない。

最後に、客観テストと学習者の自己評価の相関について、聴解と読解の2つのスキルだけではあるが、調べてみた。結果は表1の右側に示すとおりである。4-2でも述べたように、教師の評価は全被験者66名に対して行われたが、客観テストについては連絡の行き違いから66名中31名しかデータを得られなかった。そこで、この31名について自己評価と教師評価との相関を改めて計算しなおし、教師評価との相関結果と比べてみた。結果は表3に示すとおりである。

表3 自己評価と客観テストの結果との相関

	聴 解	読 解
客観テスト	0.075	0.360*
教師評価	0.037	0.232

df = 29      \*2.5%水準で有意

聴解では、教師評価との相関 $r=0.037$ に対し、客観テストとの相関は $r=0.075$ であり、読解では教師評価との相関が $r=0.232$ 、客観テストとの相関が $r=0.360$ であった。これを見ると、聴解の場合には客観テストとの相関と教師評価との相関との間にあまり大きな差はないようであるが、読解の場合には教師評価との相関が客観テストとの相関に比べてやや低くなっている。これは、聴解に比べて読解の方が教師評価と客観テストとの相関が低かったことに原因があると思われるが、一方で教師にとって最も評価しにくいスキルは読解であるとするBlueの見解を裏付けているようにも思える。ひとつ興味深いのは、表2で全被験者66名に対して教師評価との相関を算出した場合には、聴解で $r=0.381$ 、読解で $r=0.309$ であったものが、客観テストを実施した31名の被験者に対して教師評価との相関を算出してみると、聴解で $r=0.037$ 、読解で $r=0.232$ と、数値がかなり低くなっていることである。この傾向は特に聴解において顕著である。調査方法が全く同じであるにもかかわらずこのような結果が生じた以上、この原因は客観テストを実施した31名の被験者にその原因があったと考えるのが妥当であろう。とすれば、学習者側にある何らかの要因が、自己評価の信頼性に影響を与えているのかもしれない。次項では、自己評価の信頼性に影響を与える学習者側の要因について分析を試みる。

## 5-2 自己評価の信頼性に影響を与える学習者側の要因

### 5-2-1 学習者の日本での滞在歴

前節で、聴解力に関して66名の被験者全員について教師評価と学習者の自己評価との相関を調べた結果、その数値は $r=0.381$ であったにもかかわらず、客観テストを実施した31名の被験者について同じように教師評価と自己評価の相関を調べた結果は、 $r=0.037$ と極端に低くなったことがわかった。そこで、この31名の被験者群に自己評価の信頼性を低める何らかの要因が存在しているのではないかと推論を立て、31名の被験者データを改めて調べてみたところ、31名中29名が日本での滞在歴3ヵ月から9ヵ月と比較的滞在期間の短い者たちであることがわかった。そこでまず、学習者の日本での滞在歴の長さが自己評価の信頼性にどのような影響を与えるかについて分析してみることにした。

今回の調査では、被験者66名を日本での滞在歴6ヵ月から9ヵ月のグループ(93年10月生)と、日本での滞在歴14ヵ月から16ヵ月のグループ(93年4月生)に分けて、自己評価と教

師評価との相関を算出してみた。結果は表4と表5に示すとおりである。

表4を見るかぎり、日本での滞在期間が長くなれば、それだけ教師評価との相関も高くなっているように見える。しかし、両グループの被験者数が違うので、これを単純に結論付けるわけにはいかない。そこで、算出された相関係数の有意性検定を行い、その有意水準を比べることによって、両者の数値にどの程度の違いがあるかを見る目安とすることにした。その結果、滞在歴6ヵ月から9ヵ月のグループの相関係数 $r=0.386$ は2.5%水準で有意であったのに対し、日本での滞在歴14ヵ月から16ヵ月のグループの相関係数 $r=0.444$ は1%水準で有意であった<sup>(23)</sup>。滞在歴14ヵ月から16ヵ月のグループの方が有意水準が若干厳しいので、それだけ教師評価との相関も強いと言える。

表4 自己評価と教師評価の相関(滞在歴別)

	滞在歴(6-9)	滞在歴(14-16)
会 話	0.278	0.220
聴 解	0.206	0.445***
作 文	0.364*	0.449***
読 解	0.271	0.382**
総 合	0.386**	0.444***

滞在歴(6-9ヵ月) df=26 \*5%水準で有意  
 滞在歴(14-16ヵ月) df=33 \*\*2.5%水準で有意  
 \*\*\*1%水準で有意

表5 自己評価と教師評価の一致度と一致率  
(滞在歴別)

	一致率(%)	一致度(P)
会 話	14.3/17.1	0.71/0.70
聴 解	28.6/34.3	0.68/0.60
作 文	21.4/25.7	0.66/0.79
読 解	10.7/31.4	0.82/0.63

(滞在6-9/14-16)

次に、各スキルについても個別に有意水準を比較してみると、会話を除く3つのスキルで滞在期間の長いグループの方が相関が高まっている。中でも顕著なのが聴解で、滞在歴6ヵ月から9ヵ月のグループがまったく有意性のない自己評価をしているのに対し、滞在歴14ヵ月から16ヵ月のグループでは1%水準で有意と有意水準が極めて高くなっている。これは先に紹介した「モニターはインプットにさらされることによって開発される」という Krashen の主張が、モニター(m)に対しても当てはまることを裏付けていると言える。また、微妙な差ではあるが、4技能の中で最も多くのインプットを受けているはずの聴解が教師評価との相関が最も高かったことや、両グループの差が最も大きかったスキルがやはり聴解であったことは、自己評価と学習者の受けるインプットやインターアクションとの間に何らかの関係があることを感じさせる。

しかし、その一方で、表5の一致度について両グループの数値をtテストを使って5%水準(両側検定)で検定してみたところ、両グループの間に有意な差はないことがわかった。

以上の結果から、滞在歴が長くなるにつれて、即ち学習者の受けるインプットや学習者の経験するインターアクションの量が増えるにつれて、自己評価の信頼性にわずかながら影響が生じることがわかったが、それは決して大きなものではなく、6ヵ月という滞在期間の差は自己評価の信頼性にあまり大きな影響を与えていないと結論付けられる。

### 5-2-2 学習者の受けた学校教育のレベル

次に、学習者の受けた学校教育のレベルと自己評価の信頼性との関係について調べてみる。学習者の受けた学校教育のレベルと自己評価の信頼性との関係について述べた先行研究は特にないが、より高度な教育を受けた者の方が自己の学習をより客観的に判断することができ

るのではないかとと思われる。それゆえ、学習者の教育レベルが自己評価の信頼性に何らかの影響を与える可能性は十分にあると考えられる。ここでは、被験者を大学卒以上の学歴を持つ者のグループと、高校卒程度の学歴を持つ者のグループに分け、両者の間に自己評価の正確さについて何らかの差異があるかを調べた。結果は表6と表7に示すとおりであった。

相関係数で見ると、高校卒グループに比べて大学卒グループの方がかなり信頼性の高い自己評価ができてるように見える。また、両グループの差は滞在歴の差の場合以上に顕著であるようにも見える。しかし、有意性検定によって両グループの有意水準を比較してみると、4技能トータルの場合には高校卒グループの場合も大学卒グループの場合もともに1%水準で有意であり、両グループの間に差は見られなかった。

表6 自己評価と教師評価の相関(教育レベル別)

	高校卒業	大学卒業
会 話	0.194	0.496*
聴 解	0.385**	0.690***
作 文	0.392**	0.413
読 解	0.281*	0.400
総 合	0.400**	0.612**

高校卒業 df=37      \*\*\*0.5%水準で有意  
 大学卒業 df=13      \*\*1%水準で有意  
                          \*5%水準で有意

表7 自己評価と教師評価の一致度と一致率  
 (教育レベル別)

	一致率 (%)	一致度 (P)
会 話	5.1/13.3	0.72/0.73
聴 解	28.2/46.7	0.63/0.43
作 文	25.6/26.7	0.71/0.77
読 解	15.4/26.7	0.76/0.73

(高校卒/大学卒)

また、個々のスキルについて見てみると、会話や聴解などオーラルなスキルでは大学卒グループの方が信頼性が高く、読解や作文といった視覚的なスキルでは高校卒グループの方が信頼性が高いことがわかるが、この傾向を説明する有効な解釈は見当たらない。

自己評価と教師評価の一致度についても分析を試みた。大学卒グループの方が数値が低く、教師評価との誤差が小さいように思われるが、tテストを使って5%水準(両側検定)で検定したところ、やはり両グループの間に有意な差は見られなかった。

これらの結果から判断して、学習者の受けた学校教育のレベルの違いが自己評価の信頼性に特に影響を与えているようには思われない。

### 5-2-3 日本語の熟達レベル

最後に、学習者の日本語の熟達レベルと自己評価の信頼性との間に有意な結びつきがあるかどうか検証する。一般に、学習が進めば進むほど目標言語に関する知識が増え、それだけ様々な角度から自分の第二言語能力について客観的な判断ができるようになると思われるが、先行研究の調査結果は必ずしもそのような仮説を支持してはいない。

Janssen-van Dichten (1989) は、スウェーデンでの von Elk (1982) らの研究を参考にオランダにおいて成人学習者用の自己評価テストを開発し、同時に学習者の自己評価に関する分析を行っている。それによると、全被験者 637 名中かなり信頼性の高い自己評価をした 113 名の被験者について調べてみたところ、その内訳は上位グループ<sup>(24)</sup> 61 名に対し下位グループ 52 名で、学習者の熟達レベルと自己評価の信頼性の間に何らかの関わりがあるようには思えないと述べている。また、277 項目中被験者たちが比較的信頼性の高い自己評価をした 96 の質問項目を選びだし、正しい自己判断をした学習者たちについて分析したところ、上位グルー

プが大半を占めていたわけでもなく、特定の地域の出身者が大半を占めていたわけでもなかったと報告している。Janssen-van Dienen は以上の結果から自己評価の信頼性と熟達レベルとの間には、特に強い結び付きはないと述べている<sup>(29)</sup>。

では、今回の調査ではどのような結果が得られたであろうか。まず、教師の評価に基づいて66名の被験者を上位グループ(教師評価4.0以上)、中位グループ(同2.5~3.5)、下位グループ(同2.0以下)の3つのグループに分け、それぞれのグループについて自己評価と教師評価の一致度と一致率を算出した。結果は表8に示すとおりであった。

表8 自己評価と教師評価の一致度と一致率(熟達度別)

	上位グループ	中位グループ	下位グループ
会 話	0.0/1.17	21.2/0.48	14.3/0.83
聴 解	17.6/0.94	46.7/0.33	31.6/0.74
作 文	11.1/1.28	22.6/0.69	26.9/0.67
読 解	28.6/0.71	27.8/0.61	6.3/0.97

(一致率[%]/一致度[P])

表8を見るかぎり、被験者の日本語の熟達レベルと自己評価の信頼性との間に有意な関係があるようには見えない。むしろ、4技能中3つのスキルで上位グループの自己評価が最も教師評価とのずれが大きく、上位グループほど信頼性の高い自己評価ができないとさえ言えそうな結果である。しかし、今回と同じ被験者を対象とした小山(1995a, 1995b)の調査で、学習者の自己評価がレベル3に極端に集中する傾向が見られたことを考えれば、これはある意味で当然の結果かもしれない。というのも、上位グループも下位グループも関係なく多くの学習者がレベル3を選択していれば、上位グループや下位グループほど誤差が大きくなり、一致率の数値が悪くなるのは当然だからである。

また、上位グループほど自己評価と教師評価との誤差が小さく、下位グループほど教師評価との誤差が大きいと言えるかどうかを、別の角度から分析するために、教師評価と自己評価との誤差を教師の評価結果とを比較し、その相関を調べてみた(表9)。

図9 自己評価・教師評価の誤差と学習者の熟達度との相関

会 話	聴 解	作 文	読 解	総 合
0.169	0.124	0.206*	-0.197	0.070

df=64 \*5%水準で有意

結果は、両者の相関は極めて低く、やはり両者の結び付きを証明するようなものではなかった。以上の結果から、学習者の日本語の熟達レベルは、自己評価の信頼性に有意な影響を与えていないと結論できる。

## 6 調査結果の分析と考察 (Discussion)

本調査で分析を行なった3つの要因の中で、日本での滞在歴に関してはややその影響が見られたものの、その他の2つの要因については自己評価の信頼性に特に影響を与えないことが判明したが、ここではその理由について「メタ認知ストラテジー」の観点から検討を加えたい。

「メタ認知」というのは心理学の世界で使われている比較的新しい用語であるが、実のところまだ衆目の一致する定義があるわけではない。波多野(1980)の言葉を借りれば、「自らの思考・知識の状態についてわかること」<sup>(26)</sup>、もう少し噛み砕いて言えば「自分が『わからない』ことがわかること、自分が『わかる』ことがわかること・感じること」<sup>(27)</sup>ということになる。このメタ認知は学習において非常に重要な役割を果たしている。例えば、教室の中で比較的良好に起こることであるが、こちら(教師)が「ここは単に答えがわかるだけでなく、どうしてそうなるのか、その理屈をわかって欲しい」と思って一生懸命説明しているにもかかわらず、学習者が答えが合ったという事実のみに満足してしまっている場合、学習者にはこちらの意図がわかってもらえず、かえって煩わしげな顔をされてしまったなどという経験をしたことはないだろうか。この場合、その学習者は自分がまだ本当にはわかっていないということに気がついておらず、メタ認知が適切に機能していない状態にある。つまり、自分がわかっていないことを自覚できること、自分がわかっていることを自覚できることというのは非常に重要なことで、これができなければ自己の学習を修正したり正しい情報を求めたりするのが遅れ、理解そのものの質に影響を与えかねないばかりでなく、場合によってはより深い理解を求めようとする動機づけをも失わせてしまうことになりかねないのである。

本論文で研究テーマとして取り上げてきた「自己評価」も最も重要なメタ認知ストラテジーのひとつである。O'Malley et al (1985) は、言語学習におけるメタ認知ストラテジーを以下の9つに分類している。

- 計画化(Planning) : 1. 先行オーガナイザー、2. 目標や目的を立てる、3. 注意する、4. 自己管理、5. 先行準備、6. 話すことを遅らせ聞くことだけにする
- モニタリング(Monitoring) : 7. 自己モニター
- 評価(Evaluation) : 8. 自己評価、9. 自己強化

ところで、このメタ認知ストラテジーについて O'Malley et al (1985) による非常に興味深い調査報告がある。O'Malley らは、アメリカの高校の ESL の学生 70 人とその担当教師 22 人に対して、学習者が教室内外の言語学習活動でどのような学習ストラテジーを使っているかについてインタビューを行ない、初級者と中級者との間に学習ストラテジーの使用に関する違いが見られるかどうか調査を行なっている。その結果、認知ストラテジーとメタ認知ストラテジーの使用について2つの興味深いことが判明した。

そのひとつは、学習者は初級者も中級者も関係なく総じて認知ストラテジーはよく使うが、メタ認知ストラテジーはあまり活用せず(全体の30%程度)、その重要性についてあまりよく認識していないということである。中でもモニタリングや評価のストラテジーの使用頻度は極めて低く、二つ合わせてメタ認知ストラテジー全体の15%から20%を占めているにすぎないことが判明した<sup>(28)</sup>。

これまで、言語学習におけるメタ認知ストラテジーの重要性についてはいろいろ述べられてきた。しかし、ストラテジーの研究はこれまでひとつひとつのストラテジーを記述し、分類することに精力を傾けてきたが、個々のストラテジーがどの程度の頻度で使われているかについてはあまり調査されてこなかったのではないだろうか。もし、O'Malley らの調査結果が一般化できるようなものであるとするなら、学習者がより信頼性の高い自己評価をすることができるようになるためには、まず学習者にモニタリングや評価のストラテジーの重要性

に気づかせ、「意識化」させることから始めなければならないということになるだろう。

しかし、それ以上に興味深いのはもうひとつの点である。O'Malley らによれば、自然な言語運用場面における正確さや適切さをチェックする自己モニター(本論文におけるモニター(m))は、初級者より中級者のほうがやや使用頻度が上がり、学習が進むに連れて徐々に活用できるようになるが、言語学習の成果をチェックする自己評価のストラテジーは学習が進んでも使用頻度はまったく変わらないとのことであった<sup>(29)</sup>。本論文では、ここまでモニター(m)と自己評価を別物として扱わず、自然な言語運用場面で行なわれるモニター(m)活動の結果、より信頼性の高い自己評価ができるようになると考えてきた。しかし、この調査結果は、モニター(m)による「インフォーマルな自己評価」が自己評価フォームの質問に答える「フォーマルな自己評価」に必然的に転化するわけではなく、むしろインフォーマルな自己評価とフォーマルな自己評価はまったく別の認知プロセスである可能性さえ示唆したと言ってよいのではないだろうか。少なくとも、フォーマルな自己評価がインフォーマルな自己評価の上に成り立っていることは確かであろうが、インフォーマルな自己評価がフォーマルな自己評価に転化するには、「意識化」のプロセスを経なければならないということであろう。今回の調査で得られた学習者の自己評価と教師評価・客観テストとの相関は、必ずしも満足できるほど高いものではなかったが、その原因は学習者の持つバックグラウンド、つまり学習者の学歴や第二言語の熟達度などといったものにあるのではなく、これまで度々指摘されてきたように、自己評価や自律学習の訓練不足によるものである可能性が強いと思われる。本論文で調査した学習者の4つの内的要因がいずれも自己評価の信頼性に有意な影響を与えなかったのは、その意味で当然の結果だったのかもしれない。今回の調査は改めて自己評価に関する学習者トレーニングの必要性を露呈したと言ってよいのではないだろうか。

## 7 おわりに (Conclusion)

最後に今後自己評価に関する研究を進めていく上で、課題になりそうなことを今回の調査に対する反省も込めて述べておきたい。ひとつは、学習者のための自己評価トレーニングに関することである。既に述べたように、日常の言語運用場面で自分の言語運用をチェックするということと、自分の学習の成果を自己評価するということはどうやら別物のようである。とすれば、学習者が自分の学習の成果や進歩の度合いを客観的に判断することができるようにするために、自己評価フォームとは別の何らかの手掛かりを提供することが必要となってくる。もちろん、自己評価の質問表も学習者に自己の言語学習の現状や成果について振り返る機会を提供してくれるであろうし、自己の言語学習の当面の到達目標についても教えてくれる。しかし、自己評価フォームに記入するという行為は、それ自体やはり個人の主観に頼った評価方法であり、客観的に自己を見つめなければ見えないこともたくさんあるのではないだろうか。これはアメリカで英語を勉強したことがある友人の体験談であるが、その友人は英語の教師にある特定の発音を直すよう何度か指摘されたそうである。しかし、その友人には自分の発音が教師に修正を促されるほどひどいものだとは感じられず、特に気にせずいたのであるが、ある日何気なく自分の会話をテープに録音し聞いてみたところ、はじめて自分の発音の悪さに気づいたということである。発音の自己評価が難しいという点に関しては、先行研究は一致した見解を示しているが、発音以外にも客観化を必要とするものがたくさんあるのではないだろうか。進歩を自覚するというのもそのひとつであろう。これは、ある



意味で肥満者の感覚と似ているかもしれない。筆者自身肥満体質であるが、自分がどのくらい肥っているかを自覚することは非常に難しい。「今日は昨日に比べて0.5キロふとったな」などと実感することは至難の技で、一年前に買ったズボンがきつくてはけなかったり、何ヶ月か前に撮ったスナップ写真を最近撮った写真と見比べてみたりしなければ、自分の肥満がどれほどのものかなどなかなか実感できるものではない。それと同様に、言語学習の場合にも、学習者が自己の学習の進展具合や成果を自覚できるように学習者の発話を定期的にテープに録音しておくとか、こまめに学習の記録を残しておくとかの工夫が必要であると思われる。自己評価はあくまでも学習者に学習のプロセスを意識化させるための手段のひとつにすぎないということであろう。

もうひとつ、今回の調査方法に関する反省点について述べておきたい。今回の調査では、先行研究の調査結果と比べて、自己評価と教師評価の間にそれほど強い相関は見られなかった。これは、自己評価に対する学習者のトレーニング不足が主たる原因だと思われるが、もうひとつデータ収集の仕方にも問題があったかもしれない。相関係数では、データの抽出が特定のグループに偏っている場合に、実際の相関よりも低い数値がはじき出されることがあるが、このことは今回の調査にも当てはまるように思われる。というのも、今回の調査に参加した被験者たちは全て初級終了レベル以上の学習者であり、被験者のレベルは2.0から4.5の間に集中していたからである。実際、同じタイプCを使った調査でOskarsson (1981)が $r=0.77$ という非常に有意性の高い数値を算出しているが、Oskarssonの調査では被験者は1.0から5.0まで広範囲に渡っていた。また、巻末の相関図を見てもわかるように、4技能の中で最も散らばり具合が激しく見える作文が、聴解の次に相関が高かったことからそのことがうかがえる。もう少し広範囲にデータを集められれば、あるいは違った結果がでていたかもしれない。これは今後の課題としたい。

## 注

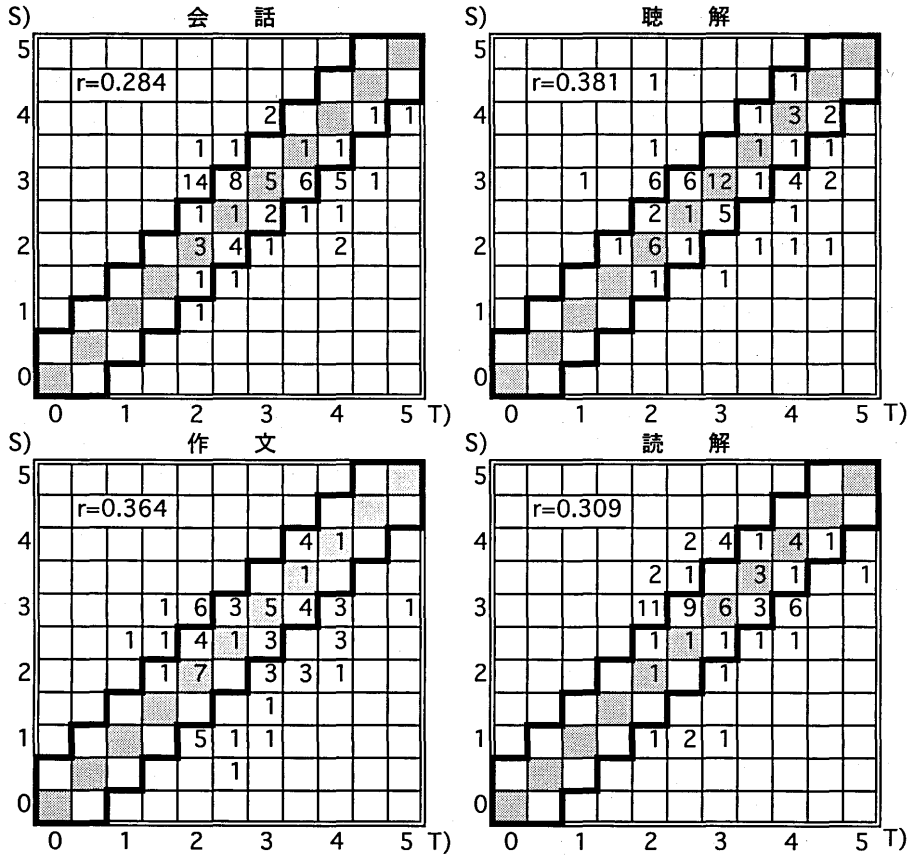
- (1) Fellenz (1976) p.257
- (2) Oskarsson (1978) pp.4-5
- (3) Backman and Palmer (1989) p.25
- (4) LeBlanc and Painchaud (1981, 1985)、Backman and Palmer (1989)、Peirce et al (1993)
- (5) Amato (1993) p.23
- (6) 吉田 (1991) p.35
- (7) Ellis (1988) p.266
- (8) Oskarsson (1978) p.25
- (9) Krashen et al (1984) pp.52-53
- (10) 『ロングマン応用言語学辞典』 p.155 "genetic epistemology" 参照
- (11) Krashen et al (1984) pp.56-58
- (12) Krashen et al (1984) p.31
- (13) Krashen et al (1984) p.33
- (14) Krashen et al (1984) p.33
- (15) Krashen et al (1984) p.34
- (16) Blanche (1988) p.82、Blanche and Merino (1989) p.325
- (17) Dornic and Ekehammar (1988) p.40
- (18) Krashen et al (1984) p.32
- (19) Backman and Palmer (1989)、Peirce et al (1993)
- (20) Krashen et al (1984) pp.34-35
- (21) 「場独立」の学習法とは、学習者が特定の項目を識別したり、その項目に焦点を合わせたりすることができ、その項目の背景となる他の項目に注意をそれされないような学習法をいう。他方、「場依存」の学習法とは、学習者が多くの項目を含む学習作業全体を見る傾向がある学習法をいう。場独立と場依存は、言語学習における認知的学習法の違いとして研究されている。  
(『ロングマン応用言語学辞典』 p.134)
- (22) Blue (1988) p.109
- (23) 本論文での相関係数の検定は全て片側検定である。
- (24) ここで言う「上位グループ」とは、調査実施前に行われたテストの成績によって分類されたものである。
- (25) Janssen-van Dielen (1989) p.41
- (26) 波多野 (1980) p.177
- (27) 波多野 (1980) p.177
- (28) O'Malley et al (1985) p.37
- (29) O'Malley et al (1985) p.38

## 参考文献

- Anderson, P. L. (1982) "Self-Esteem in the Foreign Language : A Preliminary in Vestigation." *Foreign Language Annals* 15 (2), pp.109-114
- Amato, R. (1993) 『英語教育のスタイル』(渡辺時夫 他訳) 研究社出版
- Backman and Palmer (1989) "The Construct Validation of Self-Ratings of Communicative Language Ability." *Language Testing* 6 (1), pp.14-29
- Blanche, P. (1988) "Self-Assessment of Foreign Language Skills." *RELC Journal* 19 (1), pp.75-96
- Blanche, P. and B. J. Merino (1989) "Self-Assessment of Foreign Language Skills." *Language Learning* 39 (3), pp.313-340
- Blue, G. (1988) "Self-Assessment : The Limits of Learner Independence." *ELT Documents* 131, pp.100-118
- Canale, M. and M. Swain (1980) "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing." *Applied Linguistics* 1 (1), pp.1-47
- Dornic, S. and B. Ekehammar (1988) "Note of Self-Rated Second Language Proficiency." *Perceptual and Motor Skills* 66 (1), pp.40-42
- Ellis, R. (1988) 『第二言語習得の基礎』(牧野高吉 訳) New Current International co. LTD.
- Fellenz, R. A. (1976) "Self-Evaluation : The Students." in C. Klevins (ed.) *Materials and Methods in Counting Education*, pp.256-262, New York : Klevens Publicatons, Inc.
- 波多野諠余夫 (1980) 『自己学習能力を育てる』UP選書
- Janssen-van Dieten, A. (1989) "The Development of a Test of Dutch as a Second Language : The Validity of Self-Assessment by Inexperienced Subject." *Language Testing* 6 (1), pp.30-46
- 木山美智子 (1977) 「聴解テストにおける自己評価に関する一考察」『Language Laboratory』14号 34-46 頁
- 小山 悟 (1995a) 「日本語学習者の自己評価 - 自律学習の一助として -」未公刊修士論文 南山大学
- 小山 悟 (1995b) 「日本語教育における自己評価システム導入の可能性」『南山日本語教育』第2号 南山大学
- Krashen, S. D. (1977) "The Monitor Model in Adult Second Language Performance." in M.Burt, H.Dulay, and M. Finocchiaro (ed.) *Viewpoints on English as a Second Language*, New York : Regents.
- Krashen, S. D. (1982) *Principles and Practice in Second Language Acquisition*, Oxford : Pergamon Press
- Krashen, S. D. and L. Terrell (1983) 『ナチュラル・アプローチのすすめ』(藤森和子 訳) 大修館書店
- Krashen, S. D. et al. (1984) 『第二言語の習得』(牧野高吉 訳) 弓書房
- LeBlanc, R. and G. Painchaud (1981) "Self-Assessment Scales : An Investigation into their Validity." in B. Sigurd and J. Svartvik (eds.) *AILA 81 : Proceedings* 1, pp. 194-195 (Abstract), Lund, Sweden : University of Lund.
- LeBlanc, R. and G. Painchaud (1985) "Self-Assessment as a Second Language Placement Instrument." *TESOL Quarterly* 19 (4), pp.673-687
- Morrison, D. M and G. Low (1983) "Monitoring and the Second Language Learner." in Richards and Schmidt (eds.) *Language and Communication*, pp.228-250, Longman : Longman.
- O'Malley, J. M. et al. (1985) "Learning Strategies Used by Beginning and Intermediate ESL Students." *Language Learning* 35 (1), pp.21-46
- Oskarsson, M. (1978) "Approaches to Self-Assessment in Foreign Language Skills." *Pergamon*

- Oskarsson, M. (1981) " Subjective and Objective Assessment of Foreign Language Performance. " in J. A. Read (ed. ) *Directions in Language Testing*, pp.225-239, Singapore : Singapore Univ. Press
- Peirce, B. N. M. Swain and D. Hart (1993) " Self-Assessment, French Immersion, and Locus of Control." *Applied Linguistics* 14 (1), pp.25-42
- Savignon, S. J. (1983) *Communicative Competence : Theory and Classroom Practice*.
- 田中望・斎藤里美 (1993) 『日本語教育の理論と実際』 大修館書店
- Yule, G. et al. (1987) "Learners in Transition : Evidence from the Interaction of Accuracy and Self-Monitoring Skill in a Listening Task", *Language Learning* 37 (4), pp.511-521
- von Elek (1982) "A Test of Swedish as a Second Language : An Experiment in Self-Assessment." in Lee, Y. P. et al (eds) *New Directions in Language Testing*, Hong Kong, Oxford : Pergamon
- Wangsotorn, A. (1981) "Self-Assessment in English Skills by Undergraduate and Graduate Students in Thai Universities." in J. A. Read (ed. ) *Directions in Language Testing*, pp.240-260, Singapore : Singapore Univ. Press
- Wenden (1986) "What do Second-Language Learners Know about their Language Learning ? : A Second Look at Retrospective Accounts." *Applied Linguistics* 7 (2), pp.186-205
- 吉田研作 (1991) 「外国語学習とモニター」 『日本語教育』 73号 33-43頁 日本語教育学会

### A. 自己評価と教師評価の相関図



### B. 自己評価フォーム ( Oskarsson's Type C )

(Descriptive rating scales)

Instruction: Put a cross in the boxes corresponding to your level in each of the four skills (LISTENING, SPEAKING, READING, WRITING). Only one cross for each skill!

LISTENING		
<input type="checkbox"/>	I understand the language as well as a well-educated native.	5
<input type="checkbox"/>		4.5
<input type="checkbox"/>	I understand most of what is said in the language, even when said by native speakers, but have difficulty in understanding extreme dialect and slang. It is also difficult for me to understand speech in unfavourable conditions, for example through bad loudspeakers outdoors.	4
<input type="checkbox"/>		3.5
<input type="checkbox"/>	I can follow and understand the essential points concerning everyday and general things when spoken normally and clearly, but do not understand native speakers if they speak very quickly or use some slang or dialect.	3
<input type="checkbox"/>		2.5
<input type="checkbox"/>	I can follow and understand the essential points concerning everyday and general things when spoken slowly and clearly, but in the course of conversation I often have to ask for things to be repeated or made clearer. I only understand occasional words and phrases of statements made in unfavourable conditions, for example through loudspeakers outdoors.	2
<input type="checkbox"/>		1.5
<input type="checkbox"/>	I understand the meaning of simple requests, statements and questions if they are spoken slowly and clearly and if I have a chance of asking for them to be repeated. I only understand common words and phrases.	1
<input type="checkbox"/>		0.5
<input type="checkbox"/>	I do not understand the language at all.	0

## Learner's Self-Assessment of Japanese Proficiency

Satoru Koyama  
Nanzan University

### Summary

The significance of learners' attitudes toward language learning has long been known to influence their success. This paper suggests that self-assessment is an integral part of a positive attitude toward learning and reports how language learners evaluate their own proficiency.

The study looks at two main questions : 1) whether self-assessment is meaningfully correlated with teacher assessment, and 2) whether language proficiency and background variables influence correlations of self-assessment with teachers assessment or tested proficiency. The subjects were 66 students learning Japanese as a second language at a language school in Nagoya, for the purpose of going on to university or college. They were asked to assess their own levels in each of the four language skills, using descriptive rating scales ( based on Oskarsson's Type C ). The teachers used the same rating scales for their assessment of their own students. A few days later the students took a proficiency test. Self- assessment was correlated with teacher assessment as well as the total score of the proficiency test.

Results show that the correlation between self-assessment and teacher assessment was significant at the 0.5 or 1.0 level, however this is lower than other studies. On the other hand, the length of stay in Japan turned out to somewhat influence the reliability of learner's self-assessment.