

学生の自己・他者評価に関する一考察

—教師評価との比較を中心に—

A comparative analysis of self-, peer- and teacher-assessment of English oral presentations

木原直美

Abstract

Self- and peer-assessment are powerful tools that enable teachers and learners to develop an active and student-centered learning environment. This paper investigates the relationship between self-, peer- and teacher-assessment of English oral presentations by first year college students. The findings show that, in peer-assessment, students tended to give higher scores than teachers, especially in the area of pronunciation; and in self-assessment, students who were given a lower score by teachers tended to overrate themselves, while students who were rated more highly by teachers tended to underestimate their own performances. Some possible reasons for these discrepancies are considered, and a number of possible classroom approaches to addressing the assessment gap are suggested.

1. 研究背景と目的

近年の英語教育において、コミュニケーション能力の育成は重要課題だとされている。聴衆を前にした発表・スピーチ・プレゼンテーション能力養成に関しても需要が高まってきている。発表活動は、プラスであれマイナスであれ評価を受けることから、学習者の学習意欲を高め、そのことが自分自身の自信につながる。英語教育においてもスピーチ大会、劇など有効な指導形態として多くの場面で導入されている。

筆者が担当する大学英語授業においても、シャドーイング練習の目標として、2006年4月から授業内でクラスメートを聴衆とする学習成果発表の機会を設けている。受講した学生からは「刺激になる」「緊張するけど充実感がある」など肯定的な声も寄せられているが、回数を重ねるに

つれ、以下の問題点が明らかになってきた。

問題点1：聞き役である学生に、発表者の良かった点・課題点をそれぞれコメントするよう求めたところ、良い点も課題点もコメントできない学生が多数いた。

問題点2：これとは別に、「発表はどれも良くできていた」とコメントするばかりで課題点を指摘できない学生が存在した。

問題点3：発表者の学生に、自分の発表の出来について振り返りをさせたところ、教師（筆者）の評価と大きく違っているケースが複数あった。これは、発表の出来（performanceの意、以下同じ）が特に悪かった学生に顕著であった。

問題点4：発表を重ねても、教師が前回アドバイスした点について改善がみられないケースがあった。これも、発表の出来が特に悪かった学生に顕著であった。

学生の発表に関して、聴衆の学生からコメントを求めたのは、学生が相互に刺激し合える学習者主導型の学習環境を構築しようとする試みであったが、聴衆のほとんどの学生に意見がなく（問題点1）、改善を必要とする学生の発表にもアドバイスはほとんど提供されない（問題点2）状態だったので、コメントーターはクラスの中の数人と教師に限定された。

このような事実を通して、以下のような当面の解決課題が浮かび上がってきた。まず、聴衆の学生が評価者の立場に立たされる場合の、フィルターのかからない正直なコメントである。授業では、発表者を含むクラス全員の前でコメントを求めたので、厳しいコメントを言い難かったかもしれない（問題点1）。あるいは、独自の評価が大多数評価と食い違っていることを恐れて、学生は発言を控えたのかもしれない。（問題点1、2）。

第二に、（教師から見ても）うまくいっていない学生による自己評価の正確さ（accuracy）である。発表の出来がよくないことをはずかしく感じ、わざと自分の評価を高めにしたのかもしれない。しかし発表を重ねてもほとんど教師のアドバイスが反映されていなかった（問題点4）ことを考慮すると、学生は自分の出来を正確に把握できていなかったとも考えられる。

このような観点に基づき、本稿では以下の2つのリサーチ・クエスチョンを明らかにしたい。

リサーチ・クエスチョン

- (1) 学生は、他の学生の発表をどのように評価しているのか。
- (2) 学生は、自分の発表をどのように評価し

ているのか。

2. 先行研究

小山（1996: 92-93）によれば、外国語学習者による自己評価の研究は、1970年代初頭のヨーロッパにおいて注目を集め始めた。当時のヨーロッパでは移民の増加に伴い教育ニーズが多様化し、「学習者自ら考え、計画し、学習を自己管理できるようにする」（1996: 92）言語教育観が広がっていた。なかでも Council of Europe の言語教育プログラムは、こうした自律的学習を積極的に推し進め、その一環として形成的評価を教育現場に導入した。形成的評価では、学生が自己評価を日常的に行い、授業の運営に積極的に活用されるため、学生の自己評価の信頼性に関して研究の関心が向けられるようになった⁽¹⁾。

Blanche & Merino（1989）は、こうした1979年から1986年の間に報告された21の自己評価研究について分析を行った。長阪（2003: 102, 105）は、Blanche & Merino（1989）の研究結果を以下のように整理している。7つの研究が学習者の自己評価と教師評価の相関に関して調査を行っており、うち3つの研究が学生と教師評価に相関を認め、残り4つの研究では相関は認めていない。また、9つの研究が学習者の言語能力と自己評価の関係について報告を行っており、うち5つの研究において言語能力の高い学習者は過小に自己評価すること、3つの研究において言語能力の低い学習者は過大に自己評価すること、さらに4つの研究において言語能力の高い学習者はより正確に自己評価すること、が報告されている。

ここに見られるような、研究によって学生自己評価と教師評価の相関にバラツキがみられることに関して、Blanche & Merino（1989: 326）は、Krashen（1980）のモニター能力によって説明できると主張している⁽²⁾。つまり、学習者が高

い自己モニター能力を持ち、それを十分に運用できた場合には学生と教師評価の相関は高くなり、逆に自己モニター能力を活用できない場合には低くなるというものである。

論理的思考能力など、外国語学習以外の分野においても、能力の高さと自己評価の関係について Kruger & Dunning (1999) がメタ認知能力を用いて理論を展開している。ここでいうメタ認知能力とは「自己の出来を把握でき、自分がどこで正解あるいは間違えそうなのかを判断できる能力」(訳筆者、1999: 1123) であり、Chi (1982) らが使用するモニター能力と同義だとしている。Krashen (1980) のモニター能力も自己の能力を客観視できる能力のことであるから、メタ認知能力とほぼ同義であるといえないか。Kruger & Dunning (1999) の研究でも、ユーモアの理解問題、論理問題、(母語としての) 英語文法問題のテストにおいて、高得点者ほど過小評価を行い、低得点者ほど過大評価を行う傾向が確認された。低得点者の過大評価について、Kruger & Dunning (1999) は、ある分野(例えば文法知識)における能力が低い者ほどその分野におけるメタ認知能力に欠けるため、自分の試験の出来の悪さを認識することができないと説明している。

学習者は自然に自己評価を行えるようにはならないという観点から、自己評価能力の習得過程に着目した研究も出てきている。Black (2003) は、自己評価の前にまず他者評価の導入が効果的であるとし、評価項目や評価基準は学習者と共に作成すべきだと提起している。Earl (2003) は間違いを恐れる必要のない安全な環境を構築し、評価の参考となる良い/悪い具体例を大量にインプットさせることが重要であるとしている。両者とも自己評価は才能ではなく、練習によって習得できる能力だと主張している。

プレゼンテーション発表活動における他者評価の役割に関しても研究が報告されている。

Campbell (2001) はビジネス・プレゼンテーション授業(英語を母語とする学生を対象)において、学生の他者評価は教師の評価とほぼ同程度の質・内容であったが、自己評価に関しては相関が認められなかったと述べている。英語学習者を対象とした Patri (2002) の研究でも同様の結果が報告されている。両者とも自己・他者評価を実施するにあたり、明確な評価項目と評価基準を設定し、学習者に評価の訓練を施している。

以上概観してきた通り、諸研究によって自己・他者評価の正確さや信憑性の検証が行なわれているが、明確な結論がでるには至っていない。それは、対象となる学習者の評価能力に差があるため、評価の正確さを一様に結論付けることができないためだと思われる。むしろ最近の研究では、自己評価能力を養成するための方法論に研究の視点が移ってきている。

他者評価については、Campbell (2001) や Patri (2002) の研究において教師評価の代替となり得る可能性が示唆されている。しかし、それはあくまで学習者がある程度の訓練を受けていることを前提としており、筆者が直面している受講生、そして日本における多くの英語学習者においても類推される状況においては、その理論が必ずしも有効な示唆とは言い難い。訓練を受けたことのない学習者評価の実態における、現状打開としての学習環境構築の研究が必要である。

そこで本稿では、学習者主導型の学習環境構築に向けた基礎研究として、筆者が疑問を抱ききっかけとなった授業の受講生を対象に、学生相互による他者評価、自己評価の実態調査を行い、教師の評価と比較しながら学生の評価能力に関して考察を行う。また、その考察をもとに先行研究で述べられている評価能力育成の訓練についても検討し、今後の授業運営の課題点を明らかにしたい。

3. 調査方法

調査は筆者が担当する短期大学「Basic Reading」授業で実施し、一年生 24 名を対象とした。学生は小グループ（5～6 人）に分かれ、クラスメート全員の前で英語による発表を行う。題材は、台本として Penguin Graded Readers の中から O. ヘンリー原作の『Soapy's Winter Home』（Taylor, 2000）を使い、グループごとに紙芝居形式の発表を行った。単なる朗読でなく、観衆が聞いて物語展開をイメージできるような音声ドラマを発表の目標とした。すなわち、各グループは同じ台本を使って紙芝居を行うが、音声演技については各自工夫をしなければならない。

発表した学生は自分の出来について自己評価を行い、聞き役の子は発表について他者評価を行った。クラスメートに対するプレッシャーが評価に影響を与えないように、評価・コメントは口頭でなく記述式で行い、その内容は発表者には伝えられないことを強調した。

調査は 2 回にわたって実施した。1 回目は 10 段階評価で評価点をつけ、評価項目と評価基準について教師はあえて指示を出さなかった。評価者の学生には、評価点で評価をさせると同時に、その理由についても自由記述で回答させた。2 回目の調査では、筆者が 4 つの評価項目（流暢さ、声の大きさ、発音、演技力）を指定し、それぞれの項目を 5 段階評価で得点をつけた。評価基準については学生個人で判断するよう指示をした。

2 回の発表はビデオに録画し、授業終了後、3 人の英語教師（日本人教師 1 人、英語ネイティブ教師 2 人）に同様の手順で評価を行ってもらった。学生の自己・他者評価を、筆者を含めた 4 人の教師評価と比較分析して考察を行った。

4. 結果と考察

4-1. 調査 1 教師による評価

教師 4 人による学生の表現活動の評価（以下断りがない限り「教師評価」とする）は、最高評価点と最低評価点に若干相違がみられるものの互いに高い相関が見られた。 $(0.72 < r < 0.91)$ 表 1（学生は S と表記、以下同じ）の上段は教師の評価平均（評価は 10 点満点）である。

学生による他者評価

学生相互による他者評価（以下断りがない限り「他者評価」と表記）の平均（対象学生 23 人）は表 1 下段に示すとおりである。教師評価との相関は、0.94 と高くあらわれている。表 1 と表 2 を比較してみても、顕著な相違はない。これは、一見、他者評価が教師評価の代替となり得ると報告する Campbell (2001) や Patri (2002) の研究結果に一致するようにみえる。しかし、これは学生全体の平均を比較した場合であり、個人の評価の内容に注目すると教師評価との差異が明らかになってくる。

教師の評価では、各発表者に対する評価者間の評価点の差は 2 つの事例（学生 13、学生 15: 番号は通し番号、以下同じ）で 4 点差であったのを除いて 1～3 点内に留まっていた。しかし、学生による評価では、すべての発表者に対して 4 点以上の差が見られ、最も開きが大きかった学生 1、学生 5、学生 16、学生 17 に関しては 7 点の差があった。図 2 と図 3 は、教師評価と他者評価における個人レベルの評価傾向をグラフに示したものであるが、教師と比較して学生は評価点のバラつきが大きい。（横軸は発表者、縦軸は各発表者に対する評価点。図 4、図 5 についても同じ）4 人の教師の評価傾向は比較的安定しており、最高の評価点を与えた最高点者（学生 4）と最低の評価点を与えた最低点者の一人（学生 15）についても一致し

ていた。ところが、学生が評価者に立った場合の最高点者と最低点者は、評価者によって評価が分かれ、例えば被験者学生 16 についてある学生評価者は 10 点と評価し、別の学生評価者は 3 点と評価している。この 3 点はこの学生評価者の最低点である。つまり、ある学生にとって最高レベルにある発表が、別の学生にとっては一番出来の悪いものと認識されていた。このように、学生による他者評価の評価平均は、バラつきがある評価の平均であり、個人のレベルで比較すると、教師評価との差異が明らかである。よって、Campbell (2001) や Patri (2002) が提起しているような、教師評価の代替となり得る学生他者評価とは必ずしも一致しない。

また評価点の分布にも学生と教師の相違が明らかになっている。表 2 は、学生と教師による評価点の内訳を示し、それをグラフ化したものが図 1 である。他者評価では、最高レベルと考えられる 9～10 の評価点が全体の 14.4% を占めているのに対し、教師はわずか 1.1% である。教師の方が学生よりも厳しい評価をしていることがわかる。さらに詳しくみると、10 点満点を与えた教師はいなかったのに対し、他者評価では 31 の発表 (全体の 6.1%) が 10 点満点と評価されている。10 点満点をつけた学生にとって、その発表は、改善点の余地が残されていないことを意味している。それに対し、教師は 10 点満点に値する発表はなかったと評価している。この 10 点満点に対する認識のズレが、冒頭で述べた問題点 2 の状況を生み出していると考えられる。つまり、教師は今回の発表はすべて何らかの課題点を残していると認識しているから学生にコメントを求めるが、学生は自分が 10 点満点だと評価している発表に関してはコメントができないのではないのか。

では、評価者は具体的にどのような評価項目を設定して評価を行っているのだろうか。自由記述を手がかりに、教師評価と比較しながら、

学生の傾向を探ってみたい。表 3、表 4 は、評価者が最高点、最低点をつけた理由をまとめたものである。教師が最高点をつけた理由でもっとも多かったのは、「表現力がある (3 人)」であった。学生は「聞きやすい (21 人)」という回答が 23 人中 21 人、次いで「発音が良い (18 人)」「表現力があつた (12 人)」であった。教師は、発音よりも表現力を高得点の理由として挙げている。逆に、学生相互の他者評価では表現力よりも発音を重視した評価が多い。他者評価において発音を理由に 10 点満点をつけた事例が 10 もあつたのは、学生が想定する発音の最高基準が、教師が想定する最高基準よりも低いからであろう。教師が最高点者に満点をつけなかった理由として「発音にまだ問題がある」という記述もみられた。

他者評価と教師評価にギャップが生じている原因のひとつは、発音に対する明確な評価基準がないからだと思われる。今回評価にあつたネイティブ教師や日本人英語教師は、これまでに大量の英語 (発音) のインプットを行っており、最高基準の発音とはネイティブ・スピーカーの、あるいはそのレベルに限りなく近い発音であると考えられる。それに対し、ほとんどの学生のインプット量は限られているため、学生は明確な発音評価基準を持たないまま評価を行ったと考えられる。学生の回答に「聞きやすかった」という主観的な理由が一番多く見られたのも、学生自身が評価者としての評価項目・基準を明確に持っていないことをあらわしているのではないだろうか。多くの学生が発音を重要評価項目として位置づけているものの、その基準が教師の基準より低いいため、それほど出来が良くなかった発表者にも高い評価を与え、その結果、調査 1 のような過大評価傾向が生じたと言えるかもしれない。

一方最低点について、教師評価では「発音のミスが多い (5 人)」という理由が全員に共

通していた。これは、登場人物の名前を終始間違えて発音する、基本単語のアクセントの位置が違っている、長い単語を最後まで発音できないなど、他の学生にはみられなかった基本的なミスであった。いずれも、授業での注意を聞いて練習をしておけば解決できるものばかりである。また、2番目に多かった「意味の区切りでポーズを入れていない(3人)」についても、あらかじめ台本の予習をしておけば、個人の英語能力を問わず克服できるミスである。つまり、この学生が犯した発音ミスやポーズのミスは、英語能力が劣っていることからくるミスではなく、事前の準備と練習量が不足していることに起因するミスだと類推できる。教師評価者はこの点を考慮して最も低い評価を行なったのではないだろうか。例えば、評価者の一人の教師は、「練習していない」ことを最低点をつけた理由として挙げている。学生の方は「声が小さい(15人)」という理由がもっとも多かった。これはおそらく、もっとも感覚的に理解できる項目であるからだろう。次に多かった「抑揚がない(8人)」「流暢に読めていない(7人)」「発音ができていない(6人)」は教師も減点材料として挙げている。しかし、教師は、練習不足が原因で生じるとされるミスの多さに対し、より低い評点をつけている。

学生による自己評価

表5は、教師評価と学生による自己評価を示したものである。相関は0.33となり、強い相関は認められなかった。つまり、教師の評価を基準とするならば、自己評価は必ずしも正確ではなかった。これは冒頭で述べた問題点3の状況を支持している。先行研究(Kruger & Dunning, 1999)では、ある分野における能力あるいは成績が高い者ほど自分の能力や成績を過小評価し、低い者ほど自分の能力や成績を過大評価すると報告されている。調査対象の学生に

もこの傾向が認められるのかどうか、以下検討を行なう。自己評価と教師評価の差を、教師より自己評価が高い場合はプラス(+)で表記、教師より自己評価が低い場合はマイナス(-)で表記し、それを教師評価の評価の順に並べたものが表6である(無回答の学生8は除いている)。同表では、教師がもっとも低く評価した3人(学生15、学生13、学生3)は教師よりも自分の出来を高く評価(過大評価)しており、高い教師の評価を受けた学生は自分の出来を低く(過小評価)、あるいは教師と同じように(正確に)評価していることがわかる。概ね、Kruger & Dunning (1999)の調査に一致する結果となった。

4-2. 調査2

調査1の5日後に調査2を行った。22人の学生を対象とした。その内の21人については、調査1からの継続であり、学生の通し番号は調査1で用いたものと同じ番号を用いている。調査2では、新たに学生24が加わり、学生8と学生14は参加していない。調査2では、教師が指定した評価項目を(流暢さ、声の大きさ、発音、演技力)基に他者評価、自己評価を行った。各評価項目を5段階評価で得点をつけ、合計20点とした。

教師による評価

教師4人による評価は、調査1同様、最高点と最低点に若干相違がみられるものの互いに高い相関が見られた($0.70 < r < 0.90$)。表7の上段は教師評価の平均である。

学生による他者評価

学生の他者評価の平均は表1下段に示すとおりである。教師評価との相関は、0.96と調査1同様に強い。調査1では、個人レベルでは教師との評価の違いが明らかとなった。評価項目

を用いた調査2でも、各発表者に対する評価者間の評価点の開きについて試みる。教師は、各発表者を評価する際1～7点、学生の場合は5～11点の評価点の開きが見られ(調査1は10点満点、調査2は20点満点)、教師間よりも学生間の方が評価点の開きが大きい点に変わりはないものの、学生の評価点の開き度合いが調査2ではやや小さくなっている(図4、図5)。これを評価項目ごとに試みる。教師の評価点の開きがもっとも大きかったのは「流暢さ」の項目で、学生評価者による評価点の開きがもっとも大きかったのは「演技力」、次いで「流暢さ」であった。流暢さを評価する際に評価者間で評価点のバラつきが目立つのは、他の3項目に比べ流暢さが多様に解釈でき、そのため基準設定が各評価者によって違っていたからだと考えられる。「演技力」において学生の評価に開きが目立ったのも、経験の少ない英語の表現活動に対して、学生がそれぞれ独自に評価基準を設定して評価にあたったからではないかと考えられる。

次に、評価点分布について試みる。調査1では、学生より教師の方が厳しい評価をしている。表8は、調査2における評価点の内訳であり、図6はそれをグラフ化したものである。調査1同様、教師の方が厳しい評価をしていることがわかる。調査2においても、教師が20点満点をつけた発表はなかったが、学生による評価では16の発表(全体の3.5%)が20点満点であった。これを評価項目別に見たものが、表9と図7、図8、図9、図10である。すべての項目において、教師の方が学生より厳しい評価をつけているものの、特にその傾向は「発音」(図9)の項目において顕著である。調査1で学生の発音評価基準が教師の発音評価基準よりも低いために、学生は過大評価を行ないやすいのではないかと指摘したことは、ここでも図7、図8、図9、図10によって裏付けられる。

学生による自己評価

教師評価と学生による自己評価の相関は0.57と、調査1と比べるとやや高くなっている。調査1同様、自己評価と教師評価の差を教師より自己評価が高い場合はプラス(+)で表記、教師より自己評価が低い場合はマイナス(-)で表記し、それを教師評価の順に並べる。表10はその結果をまとめたものである(無回答の学生9、学生10、学生23は除いている)。教師評価が低かった学生は、自分の出来に対して教師評価より高い評価を行い、教師評価が高かった学生は教師評価よりも低い評価を行なっている。つまり、調査2でも Kruger & Dunning (1999) の調査と一致する結果となった。Kruger & Dunning はまた、能力や成績が低い者ほど自分の出来や成績に対して過大評価を行なう傾向が顕著であり、能力や成績が高い者ほど自分の出来や成績に対して過小評価を行なう傾向が顕著であると報告しているが、調査2はこの点も裏付ける結果となった。次に、これを評価項目別に見てみる。

表11、表12、表13、表14は項目ごとに、教師の評価の順に学生の自己評価を並び替えたものである。「声の大きさ」の項目では、低い評価を受けた学生による自己過大評価、高い評価を受けた学生による自己過小評価の傾向が特に顕著であった。「発音」「演技力」に関しても、いくつかの例外は見られるが、同様の傾向が見られた。

「流暢さ」に関しては、教師の評価と学生の自己評価傾向に、明確な関連は認められなかった。先述したように、この「流暢さ」の評価項目は、教師の間でも基準が明確でないことが示唆された。このことが表11の結果に影響を与えているとも類推されるが、本稿ではその点を詳述することができていない。

5. まとめと今後の課題

本稿では、2つの調査を通して学生による自己評価・他者評価を実施し、その結果を教師評価と比較して考察を行なった。その結果、リサーチ・クエスチョンに対して以下の点が明らかになった。

リサーチ・クエスチョン (1)

学生は、他の学生の発表をどのように評価しているのか。

- ① 学生は他の学生を評価する際、教師よりも高めに評価する傾向がある。評価項目別に見ると、特に「発音」においてその傾向が顕著である。
- ② 学生は、教師が満点の評価を与えていない発表に対しても満点をつけることがある。評価項目別に見ると、特に「発音」においてその傾向が顕著である。
- ③ 学生は、「聞きやすい」(23人中21人)「発音がよい」(23人中18人)と判断した発表に対して高得点をつける傾向がある。
- ④ 学生は、「声が小さい」(23人中15人)と判断した発表に対して低得点をつける傾向がある。

リサーチ・クエスチョン (2)

学生は、自分の発表をどのように評価しているのか。

- ⑤ 教師評価が低い学生は、自分の出来を教師よりも高く評価し、教師評価が高い学生は、自分の出来を教師よりも低く評価する傾向がある。「声の大きさ」「発音」「演技力」の評価項目においても、同様の傾向を示す。

①～⑤により、今回調査対象となった学生は、クラスメートの反応を気にして、意図的に好

意的な評価をしているのではない、ということがわかった。学生は自らの判断でクラスメートの発表を評価していた。しかし、教師よりも高めに評価する傾向があるため、教師が期待しているようなアドバイスや指摘を求めることには無理がある(問題点2)。更に、教師にとっては課題の残る発表であっても、学生は最高の出来(満点レベル)であったと認識しているために、コメントに躊躇する学生も出てくる(問題点1)。また、教師の評価が低い学生ほど、自分の出来を過大に評価している(問題点3)ために、教師のアドバイスや指摘は伝わりにくく、発表力の向上に結びつかない(問題点4)。

以上のことから、調査対象となったクラスの学生に関しては、Black (2003) や Earl (2003) が主張しているように、訓練や練習を実施しないまま学生に他者評価・自己評価をさせても、教師が期待している通りの評価は行われにくいということが明らかになった。したがって、学習者主導型の学習環境構築の一助として学生他者評価・自己評価を用いる場合には、いくつかの課題点を考慮しなければならないと考えられる。

学生は全体的な傾向として教師よりも低い評価基準を設定しており、教師評価基準と一致していないことが明らかになった。これは評価者としての経験不足と、評価基準を形成する具体的な事例を十分に観ていないことに起因していると考えられる。そこで学生が評価を行なう際には、良い発表、悪い発表、満点に値する発表など、具体的な事例を豊富に与えることが有効であると考えられる。また、これを教師のフィードバックを与えながら何度も繰り返すことが重要であろう。

評価項目は、学生の英語力と評価者としての能力に合致する、具体性や客観性を持たせることが必要だと思われる。学生は23人中21人が「聞きやすい」ことを高得点の理由として挙げているが、この項目を細分化することにより、

評価基準がより明確な形で設定されるのではないだろうか。さらに、本稿の調査では学生が発音を評価項目として重要視している傾向が明らかになった。教師が発音以外の評価項目を重視する場合（例えば「演技力」「声の大きさ」）は、あえて評価項目の中から発音を排除することが有効であるといえよう。

本稿では、筆者が直面した問題点について、いくつかの調査データが得られた。また、それに基づく授業運営の課題点を得ることもできた。しかし、今回の調査では、学生の評価項目・評価基準が教師と異なることを確認できたものの、学生が持っている独自の評価項目・基準については、部分的にしか明らかにすることができなかつた。また、モニター能力やメタ認知能力の概念を調査で言及することができなかつた。

今後は、学習者主導型の学習環境を構築する立場から、英語発表活動における、①学生の発表能力向上、②そのための自己・他者評価者としての能力向上、をモニター能力、メタ認知能力の概念において研究を進めていきたい。

⁽¹⁾ 形成的評価では、学習者のフィードバックを授業運営に反映させていくため、教師は授業の中で頻繁に学習者の自己評価を実施する。形成的評価と学生自己評価の関係については Davis(2005)「大学英作文クラスにおける形成的評価方法とメタ認知の発達」『長崎外大論叢』第9号、101-108頁を参照のこと。

⁽²⁾ ここでいうモニター能力とは、外国語言語能力・パフォーマンスを自己点検し、無意識に習得した言語能力であっても自分の間違いに気づいた場合などには発話・作文などのアウトプットを自己修正できる能力を指す (Blanche & Merino : 326)。

参考文献

小山悟 「自律学習促進の一助としての自己評価」『日本語教育』88号、1996年、91-103頁

長阪朱美 「学習者の英語ライティング能力の認識：エッセイの自己評価の分析」『恵泉女学園大学人文学部紀要』15号、2003年、95-111頁

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for Learning: Putting it into Practice*. New York: Open University Press.

Blanche, P. & Merino, B. J. (1989). "Self-Assessment of Foreign-Language Skills: Implications for Teachers and Researchers". *Language Learning*, 39 (3), 313-340.

Campbell, K. S., Mothersbaugh, D. L., Brammer, C. & Taylor, T. (2001). "Peer versus Self Assessment of Oral Business Presentation Performance". *Business Communication Quarterly*, 64 (3), 23-42.

Chi, M. T. H., Glaser, R. & Rees, E. (1982). "Expertise in problem solving". In Sternberg, R. (ed.) *Advances in the psychology of human intelligence*. Vol. 1. Hillsdale, NJ: Erlbaum, 17-76.

Earl, L. M. (2003). *Assessment As Learning*. Thousand Oaks, CI: Corwin Press.

Krashen, S. D. (1980). "The monitor model for adult second language performance". In Croft, K. (ed.) *Readings on English as a second language*. Cambridge, Massachusetts: Winthrop Publishers,

213-221.

Kruger, J. and Dunning, D. (1999) "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments". *Journal of Personality and Social Psychology*, 77 (6), 1121-1134.

Patri, M. (2002). "The influence of peer feedback on self- and peer-assessment of oral skills". *Language Testing*, 19 (2), 109-131.

Taylor, N. (2000). "Soapy's Winter Home". In *The Gift of the Magi and Other Stories by O. Henry*. Harlow, Essex: Pearson Education Limited, 22-25.

表1 教師評価平均と他者評価平均 (調査1)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23
教師評価平均	4.5	6.8	3.8	8.0	5.5	6.0	4.8	4.8	5.8	5.3	6.8	6.0	3.5	4.0	2.5	4.3	4.8	6.5	5.5	6.0	5.5	7.3	7.0
他者評価平均	5.5	7.0	4.4	8.9	6.0	7.0	6.3	6.5	6.5	6.5	8.1	7.6	4.2	5.3	4.0	5.5	6.2	7.8	5.9	6.7	5.6	7.9	7.7

表2 教師評価・他者評価における評価点分布 (調査1)

	教師評価		他者評価	
無記入	0	(0.0%)	52	(10.3%)
1点	1	(1.1%)	0	(0.0%)
2点	4	(4.3%)	10	(2.0%)
3点	4	(4.3%)	21	(4.2%)
4点	20	(21.7%)	40	(7.9%)
5点	20	(21.7%)	73	(14.4%)
6点	16	(17.4%)	87	(17.2%)
7点	19	(20.7%)	89	(17.6%)
8点	7	(7.6%)	61	(12.1%)
9点	1	(1.1%)	42	(8.3%)
10点	0	(0.0%)	31	(6.1%)

図1 教師評価と他者評価における評価点分布

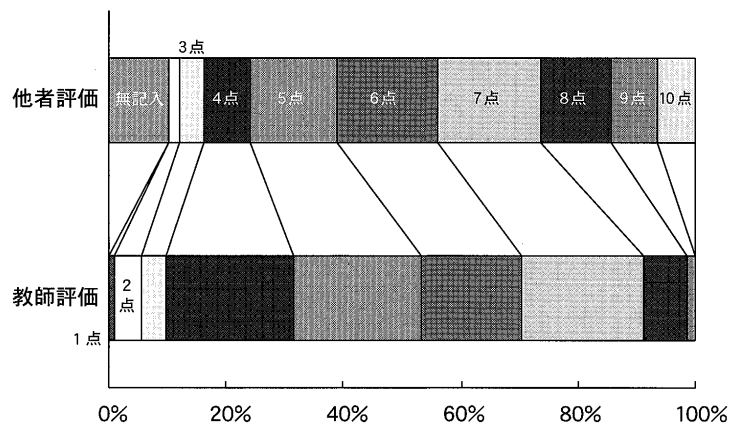


表3 最高点をつけた理由 (複数回答)

教師	学生
表現力がある人(3人)	聞きやすい(21人)
気持ちが入っている人(1人)	発音が良い(18人)
イントネーションが良い人(1人)	表現力がある(12人)
流暢に読めている人(1人)	スピードが良い(10人)
声大きい(1人)	抑揚があった(10人)
	声大きい(6人)
	流暢に読めている(7人)
	暗記している(1人)

()の人数はその評価を受けた発表者の数

表4 最低点をつけた理由 (複数回答)

教師	学生
発音のミスが多い(5人)	声が小さい(15人)
意味の区切りでポーズを入れていない(1人)	抑揚がない(8人)
練習していない(1人)	流暢に読めていない(7人)
理解していない(1人)	発音ができている(6人)
スピードが遅い(1人)	練習が足りない(3人)
	スピードが速い・遅い(1人)

()の人数はその評価を受けた発表者の数

図2 教師による評価傾向 (調査1)

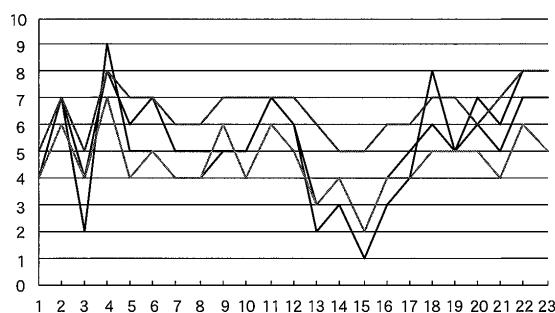


図3 学生による他者評価傾向 (調査1)

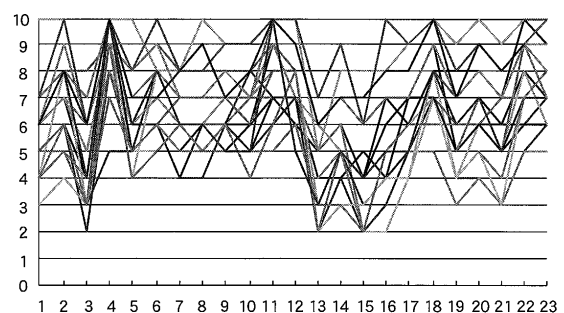


表5 教師評価と自己評価（調査1）

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23
教師評価	4.5	6.8	3.8	8.0	5.5	6.0	4.8	4.8	5.8	5.3	6.8	6.0	3.5	4.0	2.5	4.3	4.8	6.5	5.5	6.0	5.5	7.3	7.0
自己評価	3.0	4.0	8.0	8.0	3.0	7.0	6.0	-	4.0	5.0	6.0	4.0	5.0	4.0	5.0	2.0	2.0	5.0	3.0	6.0	5.0	6.0	7.0

-は無回答

表6 教師評価・自己評価間の差（調査1）

	教師評価	自己評価	差
S15	2.5	5	+2.5
S13	3.5	5	+1.5
S3	3.8	8	+4.2
S14	4.0	4	0
S16	4.3	2	-2.3
S1	4.5	3	-1.5
S7	4.8	6	+1.2
S17	4.8	2	-2.8
S10	5.3	5	-0.3
S5	5.5	3	-2.5
S19	5.5	3	-2.5
S21	5.5	5	-0.5
S9	5.8	4	-1.8
S6	6.0	7	1.0
S12	6.0	4	-2.0
S20	6.0	6	0
S18	6.5	5	-1.5
S2	6.8	4	-2.8
S11	6.8	6	-0.8
S23	7.0	7	0
S22	7.3	6	-1.3
S4	8.0	8	0

表7 教師評価平均と他者評価平均（調査2）

	S1	S2	S3	S4	S5	S6	S7	S9	S10	S11	S12	S13	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24
教師平均	7.8	13.3	9.3	17.8	12.8	13.0	11.0	10.0	11.3	17.0	14.8	5.8	7.0	9.3	9.8	15.5	12.0	14.3	8.8	16.3	15.8	10.5
他者評価	11.7	13.5	11.2	18.5	13.5	14.9	13.5	12.1	12.7	17.6	15.7	9.3	8.8	10.6	11.1	16.3	12.3	14.5	12.1	16.1	16.9	11.5

図4 教師による評価傾向（調査2）

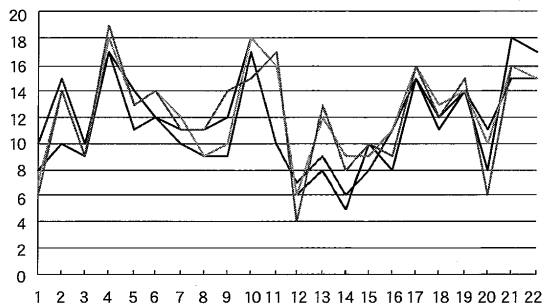


図5 学生による他者評価傾向（調査2）

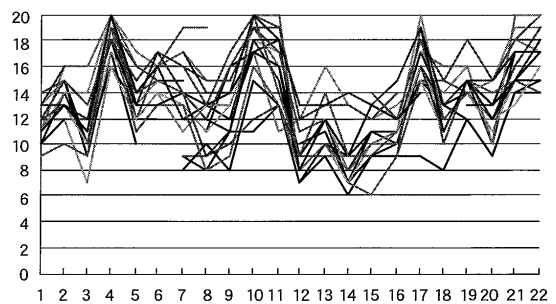


表8 教師評価・他者評価における
 評定分布 (調査2)

	教師評価		他者評価	
無記入	0	(0.0%)	64	(13.9%)
1~2点	0	(0.0%)	0	(0.0%)
3~4点	1	(1.1%)	0	(0.0%)
5~6点	6	(6.8%)	2	(0.4%)
7~8点	8	(9.1%)	22	(4.8%)
9~10点	19	(21.6%)	54	(12.7%)
11~12点	16	(18.2%)	83	(18.0%)
13~14点	13	(14.8%)	82	(17.7%)
15~16点	15	(17.0%)	77	(16.7%)
17~18点	9	(10.2%)	45	(9.7%)
19~20点	1	(1.1%)	33	(7.1%)

図6 教師評価と他者評価における評定分布 (調査2)

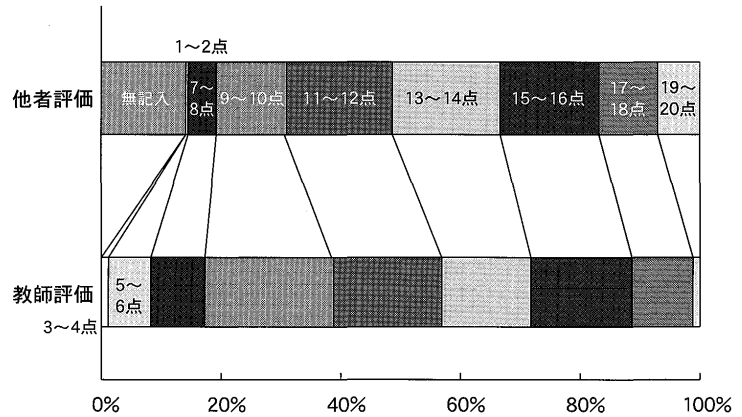


表9 教師評価・他者評価における評価項目別評価点分布 (調査2)

	流暢さ		声の大きさ		発音		演技力	
	教師評価	他者評価	教師評価	他者評価	教師評価	他者評価	教師評価	他者評価
無記入	0	54	0	47	0	48	0	45
1	5	14	11	8	4	1	9	4
2	20	87	18	62	22	56	23	92
3	31	132	25	128	43	149	30	195
4	24	113	23	139	19	150	21	84
5	8	62	11	78	0	58	5	42

図7 「流暢さ」評価点分布 (調査2)

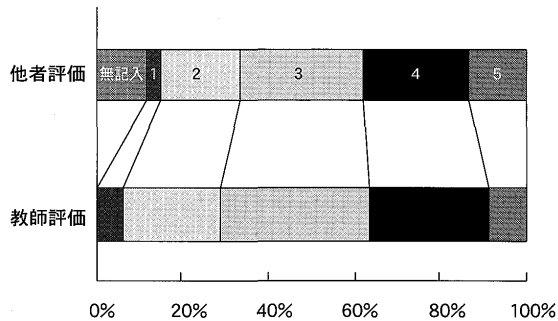


図8 「声の大きさ」評価点分布 (調査2)

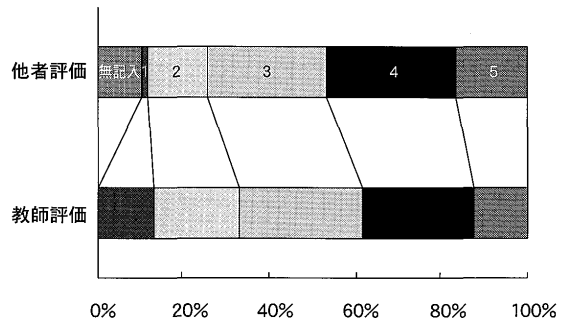


図7 「流暢さ」評価点分布 (調査2)

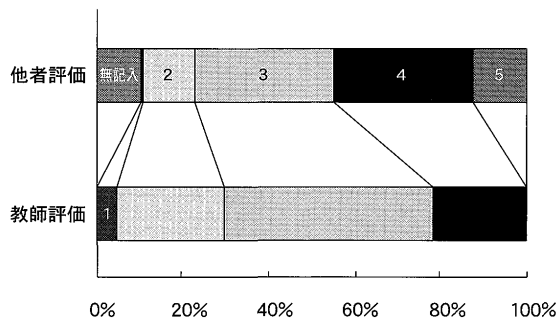


図8 「声の大きさ」評価点分布 (調査2)

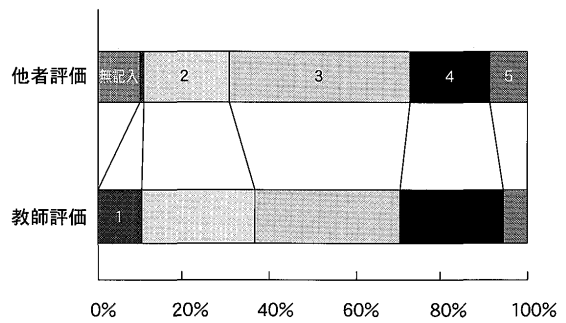


表10 教師・自己評価間の差
(調査2)

	教師評価	自己評価	差
S13	5.8	10	+4.2
S15	7.0	10	+3.0
S1	7.8	8	+0.2
S21	8.8	9	+0.2
S3	9.3	11	+1.7
S16	9.3	6	-3.3
S17	9.8	10	+0.2
S24	10.5	8	-2.5
S7	11.0	9	-2.0
S19	12.0	10	-2.0
S5	12.8	12	-0.8
S6	13.0	12	-1.0
S2	13.3	8	-5.3
S20	14.3	12	-2.3
S12	14.8	13	-1.8
S18	15.5	10	-5.5
S22	16.3	14	-2.3
S11	17.0	12	-5.0
S4	17.8	11	-6.8

表11 「流暢さ」教師評価・自己評価
間の差 (調査2)

	教師評価	自己評価	差
S15	1.3	1	-0.3
S13	1.8	2	+0.2
S3	2.3	2	-0.3
S17	2.5	2	-0.5
S1	2.5	3	+0.5
S24	2.8	2	-0.8
S16	2.8	2	-0.8
S21	2.8	3	+0.2
S2	3.0	1	-2.0
S7	3.0	1	-2.0
S19	3.0	3	0.0
S12	3.3	4	+0.7
S5	3.5	3	-0.5
S20	3.5	3	-0.5
S6	3.5	4	+0.5
S18	4.0	3	-1.0
S22	4.0	4	0.0
S4	4.8	2	-2.8
S11	4.8	3	-1.8

表12 「声の大きさ」教師評価・自己
評価間の差 (調査2)

	教師評価	自己評価	差
S13	1.0	2	+1.0
S21	1.3	2	+0.7
S1	1.5	2	+0.5
S16	1.8	2	+0.2
S17	2.0	3	+1.0
S7	2.5	3	+0.5
S24	2.5	2	-0.5
S15	2.5	3	+0.5
S2	3.3	3	-0.3
S3	3.3	3	-0.3
S19	3.3	3	-0.3
S5	3.5	3	-0.5
S6	3.5	3	-0.5
S20	3.8	3	-0.8
S22	4.3	4	-0.3
S4	4.5	3	-1.5
S11	4.5	3	-1.5
S12	4.5	3	-1.5
S18	4.8	3	-1.8

表13 「発音」教師評価・自己評価間
の差 (調査2)

	教師評価	自己評価	差
S15	1.5	3	+1.5
S3	1.8	4	+2.2
S13	2.0	3	+1.0
S16	2.3	1	-1.3
S1	2.5	2	-0.5
S24	2.5	2	-0.5
S21	2.5	2	-0.5
S6	2.8	3	+0.2
S7	2.8	3	+0.2
S17	2.8	3	+0.2
S5	3.0	3	0.0
S19	3.0	2	-1.0
S12	3.3	3	-0.3
S18	3.3	2	-1.3
S20	3.3	3	-0.3
S2	3.5	2	-1.5
S4	3.5	3	-0.5
S22	3.8	3	-0.8
S11	4.0	3	-1.0

表14 「演技力」教師評価・自己評価
間の差 (調査2)

	教師評価	自己評価	差
S13	1.0	3	+2.0
S1	1.3	1	-0.3
S15	1.8	3	+1.2
S3	2.0	2	0.0
S21	2.3	2	-0.3
S16	2.5	1	-1.5
S17	2.5	2	-0.5
S5	2.8	3	+0.2
S7	2.8	2	-0.8
S24	2.8	2	-0.8
S19	2.8	2	-0.8
S6	3.3	2	-1.3
S2	3.5	2	-1.5
S18	3.5	2	-1.5
S11	3.8	3	-0.8
S12	3.8	3	-0.8
S20	3.8	3	-0.8
S22	4.3	3	-1.3
S4	5.0	3	-2.0