



「観察」再考

根岸 雅史 Negishi Masashi
(東京外国語大学教授)

1. <参考資料>の功罪

国立教育政策研究所より、『評価規準の作成、評価方法の工夫改善のための参考資料（中学校）—評価規準、評価方法等の研究開発（報告）—』（以下、<参考資料>）が発表されたのは、観点別・絶対評価が導入される直前の昨年2月のことであった。絶対評価の導入に戸惑っていた教師には、貴重な資料となったに違いない。

<参考資料>で示されたものは、評価規準（のりじゅん）と言われるもので、そこには評価の視点が示されている。たとえば、「話すこと」の「表現の能力」についての評価規準は、「初歩的な英語を用いて、自分の考えや気持ちなどを正しく話すことができる」と「初歩的な英語を用いて、場面や相手に応じて適切に話すことができる」となっている（下線筆者）。しかしながら、どの段階ではどの程度の「正確さ」や「適切さ」を求めるのかという基準（もとじゅん）は示されておらず、これらは各学校にゆだねられている。

この<参考資料>は、後半は評価事例が示されている。英語では、ある授業の1レッスン分を取り上げ、その評価方法を具体的に示している。ところが、この評価事例が大きな混乱をもたらすことになる。毎時間の評価方法等の欄には、それぞれの活動ごとに「活動の観察」「生徒の応答」「録画チェック」と書かれている（「生徒の応答」も「（活動の）観察」により評価するものと思われるが）。これらは、1時間に2～4回行われるようになっていく。また、後日、「インタビューテスト」3回および「録画チェック」2回が行われることになっている。

これが示されたことにより、従来のペーパーテス

トによる評価がふさわしくないような生徒の特性を「観察する」という発想が広まったことはメリットとして考えてよいかもしれない。しかしながら、そのために、多くの教師は「絶対評価」＝「観察」というイメージを持ってしまったのも事実である。公開授業などの指導案を見ると、あらゆる活動の評価に「観察」と書いてあることがある。しかしながら、授業を見ても、観察を行っている様子は特に見えない。だとすれば、このような形式ばかりの評価計画を書いても、意味はないであろう。指導計画の中では、本当に観察する予定のことを書くべきであり、そのほかの指導目標に関わる点については、いつの機会にどのように評価するのかを示すべきである。

では、なぜ<参考資料>がこのような誤解を招いてしまったのか。それは、<参考資料>の評価実践が、「評価実践例」と「評価事例集」の2つの役割を担ってしまったからではないだろうか。「評価事例集」としては、たくさんの評価事例を載せることに意味があった。しかし、これを実際の授業計画に盛り込む形で、「評価実践例」が示されたために、評価の「過積載」となってしまった。しかも、ペーパーテストなどの役割についての言及がほとんどなされなかったことも混乱を招いた原因かもしれない。評価実践例としては、1レッスンの評価計画の現実的な形をペーパーテストの役割も含めて示し、参考となる評価事例については、評価計画とは別枠で、数多く例示すべきであったと思う。

2. 評価手法としての「観察」の評価

私の知るところでは、英語教育の歴史の中で、これほど評価手法としての「観察」が強調されたこと

はないのではないかと。これまでは、「評価」すなわち「テスト」であったかもしれない。しかしながら、＜参考資料＞が示されて以来、振り子は大きく「観察」に振れたのである。

ところが、この評価手法としての「観察」は、テストと比べ、その手法はまったく確立しておらず、そのためのトレーニングも行われていない。たとえば、生徒を観察する場合、どのような活動が観察可能で、どのような活動が観察不可能なのかわかっていない。また、教師は生徒のことを一度にどのくらい観察できるのかもわかっていない。さらに、いわゆる信頼性の問題もある。作文の採点やスピーキングの採点のときに、あれほど問題となる信頼性であるが、観察においてはこれがなぜかほとんど問題とされていない。何十人も生徒を観察した場合に、評価結果のぶれがないとは思えない。おそらく、今必要なのは、この観察という手法の特徴を、メリット・デメリットも含めて、冷静に評価することではないだろうか。

3. 音読（発音）の評価

このような状況を踏まえた上で、今回は音読（発音）の評価について考えてみたい。音読の評価というのは、＜参考資料＞でいうところの、「読むこと」の「表現の能力」にあたり、この中には「正確さ」と「適切さ」の規準が入っている。このうちの「正確さ」の評価を行うにあたり、「ペーパーテスト」とそれ以外の方法を比較することで、それぞれの評価方法の特徴を見てみる。

3-1. ペーパーテスト

「ペーパーテスト」による発音の評価は、その高い実用性ゆえか、いまだに多くの定期試験で用いられている。しかし、このテストは生徒に実際に発音させているわけではないために、現実の発音能力を反映していない可能性がある。最近の研究では、とりわけ、中学生などの初学者の場合は、ペーパーテストと実際の発音との関連性の薄いこと、また、強勢問題よりは音素識別問題において、その関連性の薄いことがわかってきている。では、このテストの波及効果はどうであろうか。このテストがあるため

に、生徒が実際に発音練習をするのであればいいのだが、紙の上でのテストに対しては必ずしも実際の発音練習による対応を期待できないように思われる。

3-2. 観察による評価

次に、教室活動としての音読（発音）を、観察により評価する手法について考えてみよう。この手法は、実際に生徒の発音を聞いている点において、高い妥当性が期待できる。ただ、この場合は、教師がそれぞれの発音に対して、主観的な判断を行わなければならない、この点では、信頼性の問題が出てくる可能性がある。また、一度に全員の音読を観察できないことから、観察する対象が、毎回異なってしまうという点でも、信頼性に疑問符が打たれるだろう。さらに、実用性はどうかであろうか。これはペーパーテストに比べれば、授業中の観察・記録が必要となり、はるかに手間がかかる方法であることは確かであろう。

3-3. MD・テープの提出

これは、家で音読したものを MD・テープに録音して提出させる手法である。これも、実際に発音を聞いているので、妥当性は高いと思われるが、採点における主観が入るために、信頼性は問題となる。しかしながら、教室活動の観察とは異なり、同一のものが読まれるために、この意味では、公平さが保たれている。また、この手法は、採点の手間はかかるが、実施上の（教師の）手間はかからない。おそらく、この手法の最も魅力的な点は、波及効果があるということであろう。というのは、個人で録音するために、生徒は自分で納得するものが録音できたと判断するまで練習する傾向にあるからである。

こうして見てきたように、「音読（発音）の能力」ひとつを評価するといっても、いくつかの手法が考えられ、みなメリット・デメリットがある。したがって、評価計画の作成に当たっては、それぞれの特徴をよく理解して、その中で取捨選択していくべきであろう。