

スピーキング技能の系統的伸長のための中高一貫シラバスの開発

— 評価規準・基準の設定と具体的な指導計画の策定のために必要なこと —

三浦 省五 松浦 伸和 池岡 慎 大野 誠
千菊 基司 多賀 徹哉 高森 理絵 山岡 大基
山田佳代子 幸 建志

1. はじめに

広島大学附属福山中・高等学校（以下、当校）英語科では、2002年度より中学校を対象に、2003年度より高等学校を対象に、英語科における「目標に準拠した評価」（以下「絶対評価」）に関する全国アンケート調査をそれぞれ2度実施し、新しい評価がもたらした教員の意識の変化を分析・考察し、新しい評価法の導入のねらいを達成するための課題を明らかにした（三浦ほか2003, 2004, 2005）。

その分析の結果、「表現の能力」、なかでも「話すこと」の能力伸長におけるノウハウの蓄積がすすんでいないことが浮き彫りになった。原因の一つに、アンケートの回答にも見られたが、複数の教員で生徒のパフォーマンスを評価するときの基準設定の難しさがあげられる。テストの開発自体が現場に委されているのだが、タスクの設定、実施方法の検討、採点規準・基準の設定をゼロからやるのは大変なことであると想像できる。高等学校では、まだ65%の学校が、担当学年で評価基準を決めていると回答した（三浦ほか2005）。

『英語が使える日本人』育成のための行動計画』が発表されて、発表技能の育成にもさらに関心が高まってはいると思われるが、特別な予算措置がなされた一部を除き、ほとんどの授業環境（授業時数・クラスサイズ・学習指導要領など）が大きく変わったわけではない。また、話すことを測定するためには、実際に一人一人の生徒に話す機会を与えないといけないことから、直接測定する機会を設けることが困難であることは否定できない。

しかし、テストの波及効果は誰もが認めるところである。各現場において、いかにして信頼性の高い、「話すこと」を直接評価する方法を開発するかは英語教師の責務と考えられる。

当校英語科では、4年生（高校1年生）が受講するオーラルコミュニケーションⅠのまとめの段階のテストにおいて、そのパフォーマンス分析に教員全員が取り組み、ルーブリック（評価規準・評価基準と連動した、生徒のパフォーマンスの特徴の記述）作りを試みた。この取り組みを始点とし、英語科教員の共通認識を深め、前後の学年における取り組みに系統性を持たせることにつなげなければならない。本研究の目的は、その過程と留意点をここに記し、試行的ルーブリックを提示し、「絶対評価」の理念を活かしたシラバス作りの推進を提案することである。

2. ルーブリックとその役割

ルーブリックとは、生徒のパフォーマンスを評価する際に教師が用いる評価規準・基準を一覧にまとめたものである（Evans 2004）。ルーブリックの利点は、生徒がルーブリックを見て、目標設定が明らかになるため、望ましい学習行動を生徒から引き出すことができる点である。また、シラバスを作成する際、どの技能を、いつの段階のどの授業で焦点化して扱うか、いつまでにどのレベルにまで生徒の力を高めておくか、を明示することが必要である。ルーブリックを作成することは、その第一歩となる。

ルーブリックは、全体的（holistic）と規準準拠的（criterion-based）の2つの型があり、目的に応じて使い分けるのがよい。

全体的評価のためのルーブリックでは、生徒のパフォーマンスを総括的にグレード毎に記述する。「ACTFL ガイドライン」などが典型的なものである。熟達度を見る場合など、総合的な評価に向いている。

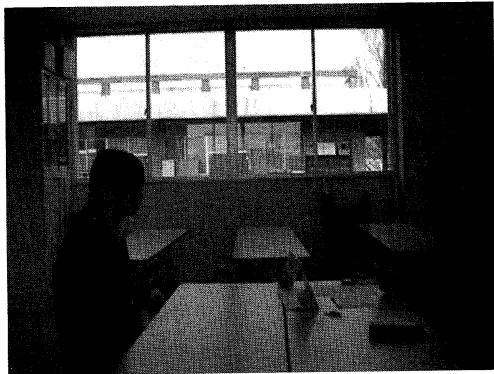
規準準拠的ルーブリックでは、観点を縦軸に、評点を横軸に格子状の表を作り生徒の行動の特徴を記述す

ることで、生徒のパフォーマンスを分析的に考察できる。したがって、形成的評価の場面でのフィードバックに適している。

3. パフォーマンス分析に用いたテストについて

2004年度オーラル・コミュニケーションⅠにおいて、インタビューテストを各学期に1回ずつ行い、その結果を表現の能力の評価に利用した。本研究においては、学年末に実施したインタビューテスト¹⁾を分析の対象とした。

テストにおいて、試験官は「親」、受験者を「子」と設定し、携帯電話を買ってもらうように2分間で交渉する。トピックはあらかじめ提示しておくが、「親」は携帯電話は不要だと考え、必ず一回は「子」に反論し、考え直すように質問する。反論は複数用意するため、「子」は「親」の発言を聞き取り、それから反論をしなければならない、というものである。



インタビューテストの様子

このテストの前に、あらかじめトピックに関連した文章を英語で読み、また、賛成・反対に分かれてディベート形式の表現活動を行うなど、練習する時間は十分に与えた。生徒はまた、テストの構造について書面で連絡を受けており、連絡から実施まで一週間あった。その際、親子間の関係は言語形式にまで反映させなくてもいいことや、「父」の性格的な問題は、とにかく反論されるということ以外は考慮しなくて良いこともあわせて連絡した。

テスト当日は、「子」が自分であらかじめ準備する部分と、「子」が「親」の対応を聞いてからその場で対応する部分に分けてパフォーマンスを測定した。

活動形式としては、オーラル・コミュニケーションⅡで扱う「討論」に近いものではあるが、設定は親子間の会話というインフォーマルな設定であること、生徒にとって身近な考えやすいトピックであること、事前にトピックに関連した授業を実施して生徒が「討論

に慣れるよう配慮したこと、また、事前に準備が可能な部分もテストに含めることで、高校1年生段階でも対応が可能だと判断した。

学年末の評価に際しては、依頼する表現を用いて適切な内容をともなった英語で発表でき、聞き取った英語に対して自分の見解を述べて反論できれば「合格」とし、発表内容の正確さ、適切さを評価し、A～Cで評価した。

4. 生徒のパフォーマンス評価とその分析によるループリック作り

4年終了時に目指すべきパフォーマンスのループリックを作るために、(1)33人分の生徒のパフォーマンスを抽出し²⁾、インタビューテスト時に撮影した映像から得た音声でCDを作成し、(2)教員8人それぞれが4段階で評価する。(3)8人のうち、5人以上が同じ評価を出したサンプルを、ベンチマーク（そのグレードの典型的なパフォーマンスサンプル）とし、なぜそのパフォーマンスはそのグレードと評価されるのかを記述する作業に進んだ。

各グレードのベンチマークの数はそれぞれ「4」=2、「3」=3、「2」=3、「1」=2で、合計10個のパフォーマンスを取り上げて、ループリック作成の基礎資料とした。それ以外のパフォーマンスは、評価者によるばらつきが大きいので、本研究では利用対象とはしなかった。

分析結果は5つの観点で記述した。4観点は、評価規準のものを利用し、さらに「発言内容」を加え、5つの規準と4段階のグレード（評点）で枠をつくった。

英語科の教員8人を4グループに分け、各グループで各グレードのサンプルのパフォーマンスを記述し、原案とした。以下は、生徒のパフォーマンスの特徴を5観点・4段階の枠組みに沿って記述したものである。

①関心・意欲・態度

評点「4」

- ・間違いをおそれず、積極的に相手の質問に答えようとしている。
- ・不自然な間をおかず、対話が続くよう努力している。相槌を打つなど、相手の話に関心を持って聞いている。

評点「3」

- ・準備したことにおいては反応がよいが、用意した答え以外の場合は答えを考える間が長くかかる。
- ・発言はききとりやすく、相手に自分の意見を伝えたいという態度が見られる。

評点「2」

- ・相手の質問の意図は理解しているようだが、考える間に黙ってしまい、自ら積極的に話そうとしているように評価できない。
- ・自らの発言に自信をもっていないようで、声も小さく、聞き取りにくい。

評点「1」

- ・応答がほとんどないため、相手の質問が理解できないのか、英語でどう表現しようか迷っているのか判断できないほどである。

②表現

評点「4」

- ・適切なスピードで正しいリズム・イントネーションを用いて、正しい発音で話している。
- ・話そうとする内容を、文法に従って正確に伝えている。

評点「3」

- ・音声面は理解に支障はなく、適切なリズム・イントネーションを用いている。
- ・質問を聞いて、自分の意見をまとめる時に長めの間が必要である。
- ・準備したことについては正確に答えられているが、その場で答える場合に文レベルで見たときに正確さに欠ける場合が数回みられる。

評点「2」

- ・音声面においては、語末の子音に不必要な母音がつくなど、日本語の影響が強く見られる。また、単語毎にブツブツ切れ、聞き手の理解に影響が出る。
- ・発言を文レベルで見たときに、正確さに欠ける場面がよく見られる。例えば動詞を用いる部分に名詞が来たりしているが、辞書で調べたような慣れない語をそのまま使っているようである。あるいは内容を伝えるために単語レベルの発言が多い例も見られ、中学段階の語彙にとどまっている。

評点「1」

- ・評点「2」よりさらに情報量が少なく、流ちょうさに欠けると判断される。

③理解

評点「4」

- ・相手の質問を正しく聞き取り、すばやく適切に対応している。

評点「3」

- ・相手の質問を正しく聞き取って対応しているが、時間がかかる。

評点「2」

- ・相手の質問を正しく聞き取っているようだが、対応

にさらに時間がかかるため、聞き取ったことを理解するのに、評点「3」の生徒より時間がかかっていると判断される。

評点「1」

- ・応答がほとんどないため、評価ができない。

④言語の知識

評点「4」

- ・自分の意見を効果的に相手に伝える表現を知っている。

評点「3」「2」「1」

- ・自分の意見を相手に伝える表現を知っている。

⑤その他： 発言内に含まれる情報

評点「4」「3」

- ・交渉するために「意見」と「根拠」を述べて相手を説得できる。

評点「2」「1」

- ・交渉するために「意見」は述べているが、「根拠」を適切に述べていない。

例えば、①～⑤の観点ですべて「3」なら評点も単純に「3」となるが、何個以上の観点で、あるいはどの観点で「3」ならそうするか、あらかじめ決めておく必要がある。

以上の結果から、「対話をする事」の到達目標のルーブリックを記述すると、各グレードを分ける特徴は以下のように集約できる。

(1) 対話を継続しようとする努力

つなぎ言葉、相槌などが見られる→評点「3」以上

(2) 音声面の特徴

日本語の影響が強く見られる発音・リズム
→評点「2」以下

(3) 文法的な正確さ

発言が文レベルで誤りが少ない→評点「3」以上

(4) 情報量

相手の質問に対して適切な返答、適切な情報構成と、複数の文からなる十分な情報。→評点「3」以上

5. カリキュラム開発のために

生徒のパフォーマンスを分析して得た情報をもとに、教師は、到達目標を意識し、系統的な技能伸長を図る活動を含めたカリキュラムの作成が次のステップとなる。4. ですでに述べた分析で明らかになった特徴(1)～(4)には、評価活動の直前に少し練習しただけではなかなか身に付かない力も含まれている。授業時間を考えると、中には科目の枠を越えて鍛えないとい

けないものもあるだろう。

例えば、教科書では複数の native speakers で交わされる会話や、non-native speakers が登場する会話でも、発言にとまどっている様子があまり記述されておらず、教科書を用いて学んでいる場合、対話を継続する努力が必要な場面にほとんど出会わない。したがって、(1)の技能は現場の教師が工夫して場面を設ける必要があることになる。

(2)(3)からは、発音・文法も文レベルで正しくコントロールできるかが評価の分かれ目になっており、初級学習者であっても、文レベルで発言させ、単語レベルにとどまることのないよう意識して指導していくのが非常に大切だと思われる。

(4)については、まずは文レベルで答えさせる習慣を授業でつけていきたい。ただし、即興の場面での話し言葉の特性を考慮し、「書くこと」で求められるようなきちんとした段落構成ではなくとも、発言全体から内容を判断することが試験官には必要である。

また、質問の意図を見抜くことが大切であり、日常の授業においても、例えば、日本の習慣を外国人に説明するタスクにおいて、どの情報を含めるべきか、それはなぜか、というところから思考させる機会を与えないと、このような技能の発展には結びつきにくいと考えられる。

学習指導要領では、「言語活動の取り扱い」、において、「指導上の配慮事項」が列挙され述べられているが、実際の授業にどのように反映させるかは現場に判断が委ねられている。高等学校の場合は、生徒の実態把握から始め、卒業時の目標に照らし合わせて生徒の能力を伸ばせるよう、どのグレードをどの学年での到達目標とするのかを検討し、シラバスを開発していく必要があるだろう。

6. 課題と展望

本研究において、分析の対象とした各サンプルが2分間であること、「対話をする」という言語活動だけであることを考えると、生徒の実態と言うには不十分な面もある。さらに、一致したサンプル数から、教員間の認識にもまだばらつきが見られることがわかった。したがって、このルーブリックは、あくまで試行的なもので、対象となったタスクのために限定される。

しかし現実には、このようなベンチマークサンプルの検討会を開くこと自体が、「学年担当者間での協議」に限定されているのが実情であることを考えると、現時点においては、この共同研究で行った活動が、学校を挙げて取り組む、実効性のあるカリキュラムの実現への現実的な第一歩であろう。ただ教科書に沿って授

業をし、単元毎に評価規準を明らかにしても、それらが学年終了時まででどのようなようにつながるのかが明らかにされ、教師間で共有されていないと、教師も生徒も評価活動に忙しいだけとなり、「絶対評価」の理念が活かされない結果となるので注意したい。

今回の分析の対象として利用したテストは、事前にテストのトピックに関連した活動を授業で行ったことと活動の性質が中・上級者向けだったことで、中学の特に入門期のカリキュラム開発のデータ収集にはそのまま利用はできないと言える。今後は、学年の枠を越えて実施できるテストを開発し、中高とも同じ評価スケールでパフォーマンス分析を行えば、系統的な技能伸長のための指導法の検討が可能になるであろう。

また、生徒に到達目標のルーブリックを公開することが、学習目標を意識した、自律した学習者を育てることにつながる、という認識を教員が共有する必要がある。たとえば生徒の自己評価も、計画的にカリキュラムに位置づけ、効果的に生徒を啓発することも可能となるだろう。

【注】

- 1) このインタビューテストは、本来、以下のような学年末の単元において、これまでの学習成果を総合的に測る目的で作られたものである。

【本課の目標】

- オーラル・コミュニケーションⅠのまとめとして、日常の「対話をする」場面において、
- (1)自分の考えを英語で伝える。
 - (2)相手の意向を理解し、積極的にコミュニケーションを図ろうとする。
 - (3)相手の質問に適切に答える。

【授業展開】

- (1)トピックに関連した題材を英語で読み、筆者の主張を理解する。
- (2)比較的平易な英語で話された、トピックに関連した意見を聞き取り、また自分の意見を述べる。
- (3)トピックについて、賛成・反対の立場に分かれ、「討論」の形式で意見交換する。

【インタビューテスト評価規準】

- (1)意見が簡潔にまとめられているか。
 - (2)質問に適切に答えられているか。
 - (3)発音・イントネーションは適切か。
- 2) サンプル数は33で、全5クラスのうち、4クラス分の生徒の中から出席番号で等間隔に抽出したものを用いた。サンプルの抽出において、男女比の差と、4年から当校に入学した生徒と、1年から在籍している

生徒との当校での学習経験の差は考慮しなかった。また1クラス分は記録が誤って消去されており、利用で

きなかつたため4クラス分を母集団とした。

【参考資料】 A sample holistic rubric

Grade (評点)	Oral Performance
4	The student answered questions with few grammar mistakes which will not interfere the listener's understanding. His or her statement had enough supportive facts which make his or her opinion persuasive. He or she had no problems understanding the spoken English given by the examiner, most of which were spoken at a normal speed. The student also effectively used fillers that showed s/he was then trying to make English sentences, or needed very short pauses to do so. His or her pronunciation, intonation, or rhythm is very good at sentence level.
3	The student answered questions with few grammar mistakes which will not interfere the listener's understanding. His or her statement is persuasive and has supportive facts. S/he had no problems understanding the spoken English given by the examiner, but s/he failed to answer promptly, and needed fairly long pauses during the conversation to think of suitable language forms to express what s/he wanted to say. But in such cases, s/he used fillers effectively and managed the time. Also, in the reaction to the refusal, some part of his or her statement show awkwardness which could have been expressed in a different and easier way.
2	Though the student tried to answer in sentences correctly, his or her reaction often had grammar errors so that the listener had to understand his or her intention by the word-level information. S/he seemed to have serious problems understanding the spoken words by the examiner, which needed to be repeated at times. Then s/he showed only incomplete sentences or word-level utterance after fairly long pauses rarely led by fillers, which often resulted in the time-management failure. Also, s/he demonstrated pronunciation problems (typically L1 interference), and the rhythm where s/he stopped almost at the end of each word.
1	The student tried to express his prepared speech but his or her problems of pronunciation, intonation, grammar or vocabulary did not get the message well understood in the interview. Particularly in the prompt reaction, he or she showed no significant supportive facts or reasoning, so that s/he just showed what he thought about the topic. It was therefore very difficult to judge whether or not s/he understood the question. Such students had difficulty in controlling the conversation in this task, though they were supposed to.

*テストの評価に使用しやすいように、テスト中に生徒が求められる発言内容の順序にしたがって、その言語的特徴を記述している。したがって、この記述の順に観点の重み付けがなされているわけではない。

参考文献

Council of Europe 2001 *Common European Framework of Reference for Languages: Learning, teaching, assessment* Cambridge University Press
 Evans, Doug 2004 Rubrics in the EFL classroom *The elt Journal* Vol. 5, No.1
 本多敏幸 2003 『中学校英語科 到達目標に向けての指導と評価』 教育出版
 Luoma, Sari 2004 *Assessing Speaking* Cambridge University Press
 松沢伸二 2002 『英語教師のための新しい評価法』 大修館書店

三浦省五ほか 2003 「英語科における「絶対評価」の現状と課題—全国のアンケート調査から—」『広島大学学部・附属学校共同研究紀要』第31号 pp.249-255
 三浦省五ほか 2004 「英語科における「絶対評価」の現状と課題(2)—全国のアンケート調査から—」『広島大学学部・附属学校共同研究紀要』第32号 pp.249-255
 三浦省五ほか 2005 「英語科における「絶対評価」の現状と課題(3)—全国のアンケート調査から—」『広島大学学部・附属学校共同研究紀要』第33号 pp.177-181