

4 技能評価に関する諸問題

山森 光陽

(国立教育政策研究所)

はじめに

本稿では、これまで外国語教育の内容、そして評価の基本的な枠組みとして用いられてきた4技能、すなわち言語能力を聞くこと、話すこと、読むこと、書くことの4つに分けて論じることの問題点を指摘する。また、学校教育における外国語教育の評価と訓練としての外国語教育の評価の違いを指摘するとともに、「生きる力」を構成する4観点、すなわち関心・意欲・態度、思考・判断、技能・表現、知識・理解で現在の学校教育において観点別評価が行われているにもかかわらず、言語系教科(国語、外国語)の観点別評価がいわゆる4技能の評価にとどまっている現状に対して考察を加える。

以上大きく分けて2つの考察を行うことによって、学校教育における外国語教育の評価と訓練としての外国語教育の評価の両者において、従来の4技能評価からの脱却の必要とその可能性について展望する。

1 4技能評価と外国語能力

4技能、すなわち言語能力を聞くこと、話すこと、読むこと、書くことといった技能の区分は、これまで外国語教育の内容、そして評価の基本的な枠組みとして用いられてきた。言語の技能をこれら4つに分類することは、現行の学習指導要領における外国語科の内容の区分においてももちろんのこと、国語においても用いられており、言語教育の内容等をこのように分類を行うことは一見したところごく当たり前の分類のように思える。

しかし、言語能力を聞くこと、話すこと、読むこと、書くことといった4つの技能に分類することは有史以来の言語教育において一貫して行われてきたわけではなく、20世紀初頭におこった教育測定運動や構造言語学の影響を受けていると考えられる。

現在我々が親しんでいる4技能に近いものを最初に提唱したのは、Lado であると考えられる。田中・青木(1985)は、Lado のテスト理論は以下の2点に要約できると指摘している。第1に、言語がコミュニケーションの習慣体型であり、我々が母国語を無意識に使えるのは、音や構造や意味のレベルでそのような習慣が身につけているからであるということである。第2に、外国語を知っているかどうかテストする際に、特定の場面でその言葉を使って何かを伝達できる、または理解できるかどうかをテストするという「場面」アプローチは、特定の場面で機能できるからと言って必ずしもその外国語を知っているとは限らないこと、そして4技能も目的に応じて個別にテストする必要があるということである。さらにLadoは言語要素を個別にテストする必要をも指摘している。

教育測定運動とはどのような運動であったかを振り返ると、ソーンダイクの「存在する

物はすべて量的に存在する。量的に存在する物はこれを測定する事ができる」というキャッチフレーズに端的に示されているとおり、学力のような構成概念をできるだけ科学的に厳密に、言い換えれば妥当性と信頼性の高い方法で測定することを目指していたのだ。

Lado が主張するように言語要素と言語技能を分類し、その分類に沿って測定を行うといった部分測定は、細分化された内容ないしは能力とテスト項目を一対一対応させることに近づくため、測定という側面から見た場合には理想的な方法であるといえるし、現在でもテストの作成の際の原則のひとつであるともいえよう。この考え方は、細分化された項目を積み上げることによって全体像に迫ることを目指していると考えられる。

しかし、学力論争に常に揺り戻しが伴うのと同様に、言語能力を統合的にとらえるべきであるとした主張もなされるようになった。Carroll (1961; 田中・青木, 1985) は、個々の構造や語彙よりはコミュニケーションが出来たかどうかをとらえる統合的なテストの必要を主張した。言い換えれば、Lado が批判した、特定の場面でその言葉を使って何かを伝達できる、または理解できるかどうかこそテストでとらえるべき対象であるとする考え方であるといえよう。

翻って日本の外国語教育において、言語の技能はどのように分類されてきたのかを検討したい。日本の小中学校、高校教育においてはどのような内容を指導すべきかを提示した「学習指導要領」によって実際の教育が行われるとともに、その成果は指導要録に記載することとされている。指導要録に各教科の評価結果を記載する際には「評価の観点」に沿って評価が行われる。そのため、指導要録に定める観点を検討することによって、日本の学校教育における外国語教育で、言語の技能をどのように分類してきたかを知ることができる。表1は戦後から現在に至るまでの中学校外国語科の評価の観点の変遷を示したものである。この表から分かるように、戦後初めて定められた指導要録においては「理解しながら読む能力」「話す技能」「書くことによって自己を表現する能力」「話された言語を理解する技能」といったように、「どのような問題を解決するために用いられる能力か」といった記述がなされている。昭和22年の学習指導要領の試案では、かなりのページ数がデュイに代表されるようなプラグマティズム的な学習観の説明についての記述がなされているが、学習指導要領の試案のもととなった考え方が評価の観点にも反映されたといえよう。

表1 指導要録における外国語科の評価の観点の変遷

昭和24年 (1949)	昭和30年 (1955)	昭和36年 (1961)	昭和46年 (1971)	昭和55年 (1980)	平成3年 (1991)	平成13年 (2001)
<ul style="list-style-type: none"> ・理解しながら読む能力 ・話す技能 ・書くことによって自己を表現する能力 ・話された言語を理解する技能 	<ul style="list-style-type: none"> ・外国語への興味・関心 ・聞く、話す ・読解 ・書き方 	<ul style="list-style-type: none"> ・聞く ・話す ・読む ・書く ・外国語への関心・態度 	<ul style="list-style-type: none"> ・聞くこと ・話すこと ・読むこと ・書くこと 	<ul style="list-style-type: none"> ・聞くこと ・話すこと ・読むこと ・書くこと ・外国語への関心・態度 	<ul style="list-style-type: none"> ・コミュニケーションへの関心・意欲・態度 ・表現の能力 ・理解の能力 ・言語や文化についての知識・理解 	<ul style="list-style-type: none"> ・コミュニケーションへの関心・意欲・態度 ・表現の能力 ・理解の能力 ・言語や文化についての知識・理解

ところが昭和30年の指導要録においては、いわゆる4技能と関心・態度を加えたものが評価の観点となっている。先に検討した、4技能を目的に応じて個別にテストする必要があるとするLadoの考え方が提示された時期と関連づけて考えると、Ladoの考え方が評価の観点にも反映されているともいえよう。

しかし、1960年代に言語能力を統合的にとらえるべきであるという主張が多くなされる

ようになったにもかかわらず、以後昭和 55 年の指導要録に至るまで、評価の観点は 4 技能を中心としたものとなっている。これは、学力の測定という観点から見て都合がよかったためと考えられる。言い換えれば言語能力を 4 つの技能に分けた方がテスト問題を作りやすいということに加えて、測定されるものこそが学力であるという考え方が 30 年あまりにわたって支配的であったからであるともいえよう。

さて、中学校についていえば、平成 13 年の指導要録以前は観点別評価といわれる 5 段階評定とは切り離された関係にあったため、観点別評価の方法はそれほど問題とはならなかった。しかし平成 13 年の指導要録からは、観点別評価をもとに 5 段階評定を行うこととなったため、どのように観点別評価を行うかが必然的にクローズアップされることとなった。現在中学校、高等学校ともに外国語科の評価の観点は「コミュニケーションへの関心・意欲・態度」「表現の能力」「理解の能力」「言語や文化についての知識・理解」の 4 つとなっている。

これらの観点のうち「表現の能力」は学習指導要領でいうところの「話すこと」「書くこと」に対応し、「理解の能力」は「聞くこと」「読むこと」に対応している。先にいわゆる 4 技能とは、言語要素と言語技能を分類し、その分類に沿って測定を行うことによって科学的かつ厳密な測定を目指したものであることから分かるように、これら 4 つの技能それぞれに対応したテスト項目ないしは測度を用いて測定し評価を行うことが求められる。

ところが、たとえば「英会話テスト」などを行った場合、聞いた内容に即して応答された内容等によって評価を行う必要があるため、「聞くこと」と「話すこと」のどちらに対応しているかを明確に分けることができない。このように、4 技能のうちの 2 つ以上の技能にまたがってしまわざるを得ない評価方法であるため、観点別評価を行うための評価方法としては望ましくないという言説が一部の中学校英語教師の中に見られるようである。

コミュニケーションの実際を考えてみると、聞きながら話したり、読みながら書いたりするということが普通に行われているのであって、これら 4 技能がそれぞれ独立した形でコミュニケーションが成立しない場合のほうが多いといえよう。また能力の評価は結局のところ答案用紙に書くにしる、口頭で応答するにしる「表現」をもとめることによってしか行うことはできないため、「聞くこと」「読むこと」からなる「理解の能力」と、「話すこと」「書くこと」からなる「表現の能力」を厳密に分けて評価活動を一貫して行うことは困難であるといえよう。

2 新しい評価の枠組み

この 5 年あまりの間に、外国語学習についての新しい評価の枠組みが提唱されるようになった。その嚆矢が Common European Framework of Reference for Languages: Learning, teaching, assessment (以下、CEF) (Council of Europe, 2002) である。この枠組みは、言語能力を一般的能力とコミュニケーション言語能力とに分けるとともに、言語活動はコンテキスト、条件、制約のもと行われ、その結果テキストを産出したり言語処理を行うが、テキストの産出や言語処理は特定の領域に属するテーマと関連していると同時に、課題の成就を目指して最も有効と考えられる方略を使うという考え方に立脚して

いる。いいかえると、これらの諸要因が密接に関連しながら言語の学習が行われるという考え方である。そして、この考え方を反映しながら6つの段階に分けた全体的な尺度を用意するとともに、各段階の「全体的尺度」を「理解すること」「話すこと」「書くこと」の3つに下位分類し、「理解すること」は「聞くこと」「読むこと」に、「話すこと」は「やりとり」「表現」のそれぞれ2つに細分化され、あわせて5つの能力分類と6つの段階からなるマトリクスによって到達目標を記述している。

CEFの注目すべき点は言語能力全体を包括した「全体的な尺度」を用意した点にあるが、実際の運用に当たってはいわゆる4技能の分類に沿いながら評価を行うように構成されており、統合的な評価にまで至っていない。しかし、この枠組みが想定している測定の実地にあたっては、直感的、質的、量的の各方法の組み合わせが必要であることを主張していることから、教育測定運動から続く厳密な測定からの脱却を試みているともいえよう。

これらの問題点を意識しつつ新たな枠組みを提案しているものとして、田中・アレン・根岸・吉田(2005)があげられる。この提案は、CEFはコミュニケーション能力を評価することを目的としているにもかかわらず、実践的使用の場の捨象と構成素の相互連関の欠如という問題を抱えていることを指摘し、通じること(intelligibility)を前提としながら、ふさわしいこと(appropriateness)、正しいこと(well-formedness)という要件を満たしつつ、実際の言語使用の場において調整可能であること(adjustability)が満たされることによって外国語を機能的に用いること(機能性: functionality)が評価の基準とならなければならないという考えにもとづいている。

以上のような4技能評価と外国語能力の見解についての変遷をたどると、言語技能を4技能に細分化した上で厳密に測定しようとする考え方のインパクトが大きく、近年に至るまで外国語教育の内容、そして評価の基本的な枠組みを規定してはきたものの、次第に真正のコミュニケーション能力を評価することが試み始められたといえよう。

3 標準テストからオーセンティックアセスメントへ

外国語学習の評価が、教育測定運動から続く厳密な測定からの脱却を試みている現状にあることを先に指摘したが、このような動きは教育評価全般についても当てはまる。

教育測定運動からの一貫した流れは、出来るだけ厳密に構成概念である学力を測定し、科学的に扱おうとしてきた営みであるといえる。しかし次第に、テストだけで学力を捉えようとしても、子どもがどこまで真に理解し、どこからが出来ないのかといった発達の真の様相をとらえることができないという問題が認識されるようになった。そして、テストに替わる「代替の評価」が模索されるようになった。

標準テストに代表されるような従来テストの問題点として、Darling-Hammond, Aness, & Falk (1995) はテストで測定される学力は限定的でありパフォーマンスを測定するには向いていないこと、知識の記憶に重点が置かれるため文脈依存によって進められる本来の学習の様相をとらえることが出来ないこと、テストで良い成績を取ることに対応することを目的とした授業に陥りやすいことなどをあげている。

一方、代替の評価においては測定(テストなどの尺度によって学習成果を量的にとらえ

ようとする)こと、査定(テストを含めて、学習者の学習成果や過程についての情報を集めること、評価(査定の結果を価値づけすること)の3つを明確に区別している。そのうえで、査定を測定によってのみ行うのではなく、ポートフォリオに代表されるように目的、計画的に学習状況についての情報を収集することによって、学習過程やその成果を丸ごととらえることが必要であると考えられるようになった。そして、学習内容を評価するのではなく、学習内容が実生活において使えるかどうかを評価しようとする「真正の評価」(authentic assessment) (Wiggins, 1989)が強調されるようになり、次第にパフォーマンス評価が注目を集めるようになってきている。

4 「生きる力」を構成する4観点と言語教育における4技能

外国語教育は様々な場面で行われているところであるが、現在のところ中学校における英語教育から多くの子どもが外国語学習を開始するという点を鑑みると、中学校における評価の枠組みを検討することは、その後につながる外国語学習における評価のあり方を考える上でも参考となるだろう。

教育課程審議会(2000)は、知識や技能の到達度はもとより、自ら学ぶ意欲や思考力、判断力、表現力などの資質や能力までを含めた学習の到達度を適切に評価する必要性を指摘し、そのための評価の観点として、4観点による評価を基本とすること適当であると答申している。つまり「生きる力(自己教育力)」を構成する要素として4観点が設定されているのである。

この点に関して、高浦(2004)は4観点について次のようにまとめている。

第1に、「関心・意欲・態度」は学ぶ力を支える情意的な側面における発達的特質を捉えようとする観点であり、一般にいわれる学習意欲、積極性、集中心、持続性、協調性、責任感などの心的傾向性を考えることができる。

第2に、「思考・判断」は、学ぶ力を構成する高次の知的機能ないしは思考技能の発達を捉えようとする観点であり、問題解決能力や問題解決の操作能力といった、メタ認知技能などの発達的な特質を捉えようとする観点である。

第3に、「技能・表現」および「知識・理解」は「関心・意欲・態度」に支えられながら「思考・判断」を働かせて学習した結果として身につけられる能力特質である。

これら「関心・意欲・態度」「思考・判断」「技能・表現」「知識・理解」といったいわゆる評価の4観点は、ほとんどすべての教科に共通の評価の観点である。ところが、外国語科の評価の観点はここで提示した能力分析的な4観点ではなく、「思考・判断」「技能・表現」の代わりに「表現の能力」「理解の能力」という内容分析的な観点が設定されている。

先に筆者は、現行の学校教育における外国語科の評価は、4つの技能それぞれに対応したテスト項目ないしは測度を用いて測定し、評価が行われている現状がみられることとともに、4技能のうち2つ以上の技能にまたがってしまうを得ない場合、それは観点別評価を行うための評価方法としては望ましくないという言説が見られることを指摘した。学習行為そのものを対象としたパフォーマンス評価の実施が求められている現在にあって、外国語科の評価の観点は未だに4技能の分類から脱却できていないことが、観点

別評価の実施そのものを困難としているのだといえよう。

また、現在の指導要録における外国語科の評価の観点に内容分析的なものが含まれることは、学校教育の目標の成就を目指す指導と評価という観点からもまた問題である。先に紹介した答申（教育課程審議会、2000）は、評価の基本的な考え方として、学習指導要領の総則では生きる力をはぐくむことを目指すべきであることが定められていることをふまえて「学力については、知識の量のみでとらえるのではなく、学習指導要領に示す基礎的・基本的な内容を確実に身に付けることはもとより、それにとどまることなく、自ら学び自ら考える力などの『生きる力』がはぐくまれているかどうかによってとらえる必要がある。」と述べられている。さらにその評価にあたっては、「自ら学ぶ意欲や思考力、判断力、表現力などの資質や能力までを含めた学習の到達度を適切に評価していくことが大切である。」とし、『『関心・意欲・態度』『思考・判断』『技能・表現』『知識・理解』の4観点による評価を基本とすることが適当である。』と提言されている。

この答申に沿って考えると、各教科における学習内容が「生きる力」を育むことを目指しており、その生きる力の下位構造として、いわゆる評価の4観点があるという構造となっていることがわかる（図1）。

ところが、先に指摘したように外国語科の評価の観点は先に述べたように能力分析的観点と内容分析的観点が混在している。すなわち、「コミュニケーションへの関心・意欲・態度」と「言語と文化についての知識・理解」については、能力分析的に観点が設定されている。一方で、「表現の能力」「理解の能力」は内容分析的に観点が設定されている。本来、学校における各教科の評価を行うに当たっては、いわゆる評価の4観点に沿って評価を行うことで「生きる力」の育ちの様相に迫ることが求められるわけであるが、外国語科の評価の観点の場合は「生きる力」の下位分類のうち2つしか設けられていないため、「生きる力」の評価を行うには不十分であるといえよう。

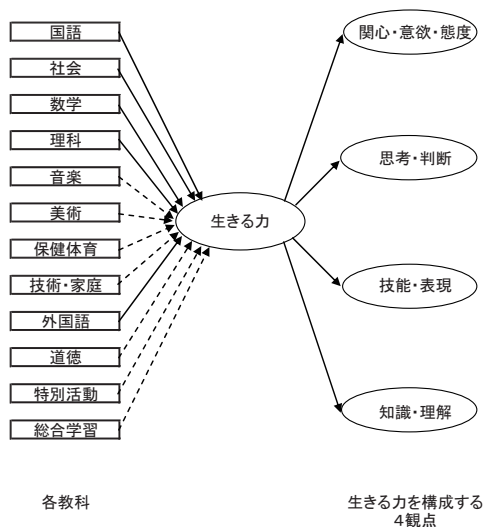


図1 生きる力と各教科との関係

5 4 技能評価からの脱却の必要とその可能性

以上のように考えると、学校教育における外国語教育であっても、また訓練としての外国語教育であっても、従来の4技能の分類に沿った評価から脱却する必要があることが理解できよう。

これまで我々がなぜ言語能力を聞くこと、話すこと、読むこと、書くことの4つに分け

て評価してきたのかを紐解くと、学習者の身につけた言語技能をより厳密に測定しようという考えが底流に流れていることを指摘した。そして4技能という分類に沿って評価を行うことは、外国語教育に対する現代的な要請と、学校教育における外国語教育に対する要請に応えられないことを論じた。

しかし、4技能という分類の起こりをたどると、それは外国語能力やコミュニケーション能力といった構成概念をできるだけ科学的に厳密に、言い換えれば妥当性と信頼性の高い方法で測定することを目指した故の帰結であるといえる。しかし現在では、真正の評価の実現のためにポートフォリオ評価などを用いたパフォーマンス評価についての研究にも精力的に取り組まれているほか、情報技術の進歩によりコミュニケーション能力を実際の場面に近い形でインタラクティブに測定することをも可能となってきた。したがって、外国語教育においても教育評価・測定研究の最新の知見と技術を導入することにより、現代的な要請に応えうる評価方法を開発し、実践に取り組まれるべきであるといえよう。

TUFS 言語モジュールは外国語「学習」に対して最新の研究成果を反映させるとともに、最新の技術を導入することによって外国語「学習」の自習システムとしての完成系に近づきつつある。しかし、教授学習過程においては常に評価がつきまとうという視点に立つと、評価の部分がやや不十分であったことは否めない。TUFS 言語モジュールが最新の情報技術を駆使しつつ、最新の研究の知見を反映させながら構成されていることを鑑みると、幸いにして教育評価・測定研究の最新の知見と技術との親和性は他の外国語学習教材と比べて極めて高い。

以上のことから、TUFS 言語モジュールの開発が21世紀COEプログラムの研究期間内で終了することなく、指導と評価が適切に一体化された新しい外国語自学パッケージとしてさらに完成型に近づくための研究が継続して推進されることを期待したい。

引用文献

- Carroll, J. B. (1961) Fundamental considerations in testing for English language proficiency, In Allen, H. B. & Campbell, R. N. (Eds.) (1972) *Teaching English as a second language: A book of readings*. McGraw-Hill. pp. 313-321.
- Council of Europe (2002) *Common European Framework of Reference for Languages: Learning, teaching, assessment*.
- Darling-Hammond, L., Ainess, J., & Falk, B. (1995) *Authentic Assessment in Action: Studies of Schools and Students at Work*. Teachers College Press.
- 北尾倫彦 (2004) 目標標準評価三年目の現状 指導と評価 50 (8), 4-8.
- 教育課程審議会 (2000) 児童生徒の学習と教育課程の実施状況の評価の在り方について (答申)
- 田中正道・青木昭六 (1985) 英語学力の本質 青木昭六 (編) 英語の評価論 大修館書店 pp. 43-69.
- 田中茂範・アレン玉井光江・根岸雅史・吉田研作 (2005) 幼児から成人まで一貫した英語教育のための枠組み リーベル出版
- 高浦勝義 (2004) 絶対評価とルーブリックの理論と実際 黎明書房
- Wiggins, G. (1989) A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70 (9), 703-713.