

「話す能力の評価に関する一考察」

広島大学大学院

萬 谷 隆 一

1 はじめに

テスト作製に際して、第一の基本的ステップは、テストする技能・能力を明確に規定することである。例えば、小学校の教師が一ケタの整数二個のたし算のテストを作製するとしよう。テストする目標技能は、 $3+2$ と $2+3$ などを一つと数えて、全部で四十五通りのたし算に正しく答えられることである。一方、英語科のテストの場合は、これほど明確に規定することは難しい。これは、生きている言葉を扱うからであろう。しかし、この問題は、良いテストを作るためには避けて通ることはできない。

小論においては、特に話す能力を測定する方法を考える基礎として、話す能力をいかに把握すべきかという問題について考察を加える。尚、対象は Proficiency test としての話す能力のテストを中心とする。

2. 言語能力に関する二つの考え方

言語能力の評価・測定分野において、大別して二つの能力観がある。一つは、Discrete Point (DP) Approach という立場、もう一つは、Integrative Approach という立場である。それぞれの立場に立って作られたテストは、DP test, Integrative test と呼ばれる。

前者においては、Lado がその中心人物である。彼の基本的な考え方は以下の如くである。

Language is built of sounds, intonation, stress, morphemes, words, and arrangements of words having meanings that are linguistic and cultural Each of these elements of language constitutes a variable that we will want to test. (Lado 1: 25)

すなわち、言語能力は個々の部分の集積であり、それら個々の部分を個別に測定することによって、最終的に言語能力全体を測ることが可能であるとする立場である。ここで言う「個々の部分」とは、音素、形態素、句、節、文や、四技能などを指す。

第二の立場は、言語の communicative use を重視する最近の傾向を反映したもので、中心的な人物は J. W. Oller である。彼は次のように述べている。

The question of language testing is not so much whether the student knows such-and-such a pattern in a manipulative or abstract sense, but rather, whether he can use it effectively in communication. (Oller 2: 198)

また、この立場の考え方の端緒となった Carroll (3: 318) は、次のように述べている。

If we limit ourselves to testing only one point at a time, more time is ordinarily allowed for reflection than would occur in a normal communication situation, no matter how rapidly the discrete items are presented. For this reason I recommend tests in which there is less attention paid to specific structure points or lexicon than to the total communicative effect of an utterance.

すなわち、言語能力は、本来個々の部分を統合的に活性化させ、意志の伝達、または受容を可能にする能力であり、この能力を部分的、個別的に測定することは不十分であるとする立場である。

以上の二者の言語能力観を、四技能能力を例にとって、単純化して図示すれば次のようになる。

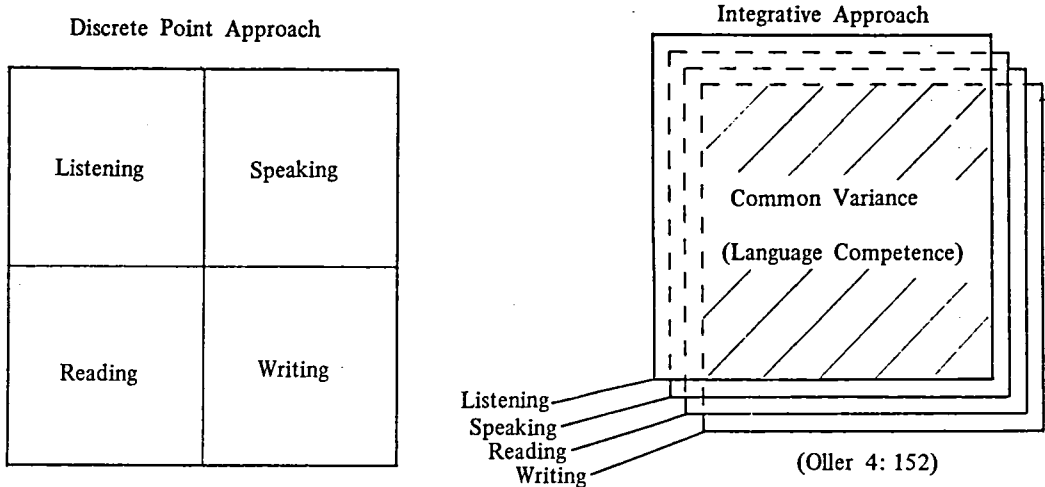


図1 言語能力に関する二つの見解

さて、両者の立場から話す能力のテストを作ると、どのようなものになるであろうか。

DP Approach の立場からは、話す能力をできるだけ細く分解して測定するテスト、例えば、文を復唱させて発音をみる音のテスト、問答で三人称単数の -s が使えるかをみる文法のテスト、絵を見せて英語の単語を言わせる語いのテストなどを使用することになる。

Integrative Approach の立場からは、習得した技能をなるべく多く使って話させ、主に意志伝達が可能かどうかの観点から採点されるタイプのテスト、例えば、インタビュー方式、スピーチ、絵の説明、などを使用することになる。

もちろん、Oller (2: 190) が、“Actually, the differences between them are not so much of type as of degree.” というように、全ての話す能力のテストを明確に上記二種類に区別は難しく、実際は個々のタイプのテストが、Integrative と Discrete Point を両極とする線上に点在していると考えるのが妥当であろう。しかし、一応は区別して、両者の立場が、話す能力を測る際に、それぞれいかなる利点・欠点があるのか論ずることは、決して無益ではないと考える。

3 話す能力のテストと Discrete Point Approach / Integrative Approach.

小論では、Davies (5: 149) の意見をとりあげ、基本的立場としたい。彼は、次のように述べている。“... the most satisfactory view of language testing, and the most useful kinds of language tests, are a combination of these two views, the analytical and the integrative.” つまり、両者の利点をうまく取り入れようという立場である。

そこで、問題は、いかに両者の利点を取り入れるかということになる。ここでは、テストにおける task と scoring の二点から考えてみたい。表1は、task と scoring をそれぞれ Integrative / Discrete Point に分けて、四つの可能性を示したものである。

①は、典型的な Linguistic Competence のテストで、音、文法、語いなどの項目中心のものである。②は、テストとして適切ではなく、例はない。③は、インタビュー、スピーチなどのテストで、採点観点が細分化されたものである。④は、③と同様で、採点が単一、総合的観点から成されるものである。

まず第一に重要であると思われるのは、task が実際の話す能力を使う場面に近いということ

である (Clark 3:38)。この点では, Discrete Point よりも Integrative の方が優れている。Task が Discrete Point であれば, 受験者が実際場面で伝達することができるかどうかは, テスト結果から間接的にしか判断することができない。Task が Integrative であれば, その判断は, より直接的になる。このことから, ①と②は, 話す能力を測る Proficiency test として不十分であると言える。」

そこで, 問題は, ③か④のどちらがより適切か, ということになる。Davies (7:220) は, ③が最も望ましいとしている。その理由は, ③の場合, task が Integrative であることによって妥当性を高め, Scoring が Discrete Point であることにより信頼性を高めているからである。このタイプのテスト例は, Carroll (8:16) の IEA International Study における Speaking test, Jakobovits & Gordon (9:54) など, かなり例は多い。

しかし, この方法には問題がないわけではない。第一に, いくつかの採点観点を定めることにより, 真に信頼性が向上するかどうか明確ではない。Francis (10:20) は, 複数の観点から採点した場合と印象による単一観点から採点した場合とでは, 両者の間に信頼性の差はみられなかったという実験結果を報告している。第二に, どういった観点を設定すべきかという問題がある。一般には, Grammar, Vocabulary, Pronunciation, Fluency といった観点が使われるが観点の種類によってテストの妥当性が左右されてくるため, どのような観点から採点すれば, 真に話す能力をとらえることができるか熟慮する必要がある。

一方, ④の方法は, 一般に指摘されるように, 採点者内で, 又, 採点者間で, 採点観点の「揺れ」があるために, intra / inter-rater reliability が低くなりやすい, という欠点を持っている。しかしながら, 言語能力が統合された動体であり, 分解して測ることができないとすれば, 単一の観点から採点することが最も望ましいであろう。信頼性を向上させるためには, 一般に, ①あらかじめ standardization を行い, 採点者の規準を一定に保つ, ②採点者を少なくとも二名以上にすることが必要である。よく言われる程には, この方法の信頼性は低くはないが, 上記のように信頼性を高めるために多大の労力と費用がかかる点が欠点となっている。

こうしてみると, ③, ④共に完全ではない。しかし, DP, Integrative 両アプローチの中道に過ぎないかもしれないが, 今後, 妥当性・信頼性を高めるために更に研究・改善の余地がある点で興味深いのは, ③であろう。

4 おわりに

小論は, 話す能力についての DP Approach ・ Integrative Approach という2つの立場を考察し, テストの task と scoring の面で両者を統合する方向性を示唆した。話す能力のテストは, 大別して, 目標能力の設定, テスト方法の決定, テストの実施という三つの問題をかかえており, そのそれぞれに困難さが伴う。今後, 尚一層の研究が望まれる。

表1 Task 及び Scoring 別にみた Speaking Testの種類

	Task	Scoring
1	DP	DP
2	DP	Integ.
3	Integ.	DP
4	Integ.	Integ.

DP = Discrete Point

Integ. = Integrative

〔参考文献〕

1. Lado, R. (1961) *Language Testing: The Construction and Use of Foreign Language Tests* (Longman).
2. Oller, J. W., Jr. (1973) "Discreted-Point Tests Versus Tests of Integrative Skills," in J. W. Oller, Jr. & J. C. Richards (eds.), *Focus on the Learner: Pragmatic Perspectives for the Language Teacher*, 184-199. (Newbury House)
3. Carroll, J. B. (1965) "Fundamental Considerations in Testing for English Language Proficiency of Foreign Students," in H. B. Allen & R. N. Campbell (eds.) *Teaching English as a Second Language*, 313-321. (McGraw-Hill)
4. Oller, J. W., Jr. (1976) "A Program for Language Testing Research," in H. D. Brown (ed.) *Papers in Second Language Acquisition, Language Learning*, 141-166.
5. Davies, A. (1978) "Language Testing," *Language Teaching and Linguistics: Abstracts*, 11, 3, 145-159.
6. Clark, J. L. D. (1979) "Direct vs. Semi-Direct Test of Speaking Ability," in E. J. Briere & F. C. Hinofotis (eds.) *Concepts in Language Testing: Some Recent Studies*, 35-49. (Teachers of English to Speakers of Other Languages).
7. Davies, A. (1978) "Language Testing - Part II," *Language Teaching and Linguistics: Abstracts*, 11, 4, 217-231.
8. Carroll, J. B. (1978) "Language Proficiency Tests Developed for the IEA International Study of Achievement in French as a Foreign Language," in B. Spolsky (ed.) *Advances in Language Testing Series: 1, Some Major Tests*, 1-48. (Center of Applied Linguistics).
9. Jakobovits, L. A. & B. Gordon (1974) *The Context of Foreign Language Teaching*. (Newbury House)
10. Francis, J. C. (1981) "The Reliability of Two Methods of Marking Oral Tests in Modern Language Examinations," *BJLT*, 19, 1, 15-23.