The Japan Association of College English Teachers (JACET)

科学的テスト時代のテストの非科学性 - テスト問題形式と配点比重についての考察

木村真治　関西学院大学
津村修志　京都外国語大学
清水裕子　立命館大学

はじめに

言語テスト研究の分野では、テストの妥当性、信頼性、実用性が主な研究対象とされるが、近年は項目応答理論を用いた能力及び項目難易度の推定の研究、また項目応答理論の応用（TOEFL、実用英語検定等）が目覚ましい進歩を遂げている。項目応答理論がテストの信頼性・実用性の向上に果たした役割は疑う余地もなく、近年では、TOEFLがコンピュータ化され、短時間に少ない項目で能力の測定が可能になった。

言語テストの信頼性・実用性に数学的理論が大きな影響を与えるようになった一方、テストの妥当性の研究は、いまだ研究者や問題作成者の経験・知識による判断に基づくところが大きい。言語テスト研究には、本質的に科学性と非科学性が混在するのであろうが、項目応答理論の急速な発展は、その科学性と非科学性の隔たりを更に顕著にしているように見える。妥当性は信頼性と並んでテストの最も重要な要件であり、妥当性の低いテストの解答データは、いくら高度に数学的な解釈をしても無意味に終わる危険性がある。

構成概念妥当性と内容妥当性

言語テストの構成概念妥当性とは、例えば、Communicative Competenceを測定するテストの作成を試みたとして、この不明朗な概念をどのようにoperationalizeするかと言う問題である。多くの場合、テスト項目の得点がこの概念を測定するものか否かが問題にされるが、本研究では、構成概念とテスト得点を直接的に捕えるのではなく、構成概念がいくつかのsub domainからなるとして、その妥当なsub domainの比率を研究対象とした。具体的には実用英語検定2級試験の筆記試験を取り上げ、その構成概念妥当性や内容妥当性を直接問題とせず、sub domainの配点比重を分析したものである。

研究の概要

本研究で取り上げる問題形式は、実用英語検定2級試験の筆記試験に出題される以下の6問題形式である。

1. 短文の語句空所補充問題
2. 和訳つき英文語句整序問題
3. 会話文整序問題
4. 長文内容一致問題
5. 長文空所補充問題
6. 長文内容一致問題

まず、実用英語検定2級試験を149名の日本人大学生に実施して項目分析を行い、この結果に基づいて、問題形式毎に正答率が中央値の項目をその形式の代表項目として選び出した。第1問題形式の「短文の語句空所補充問題」は合計30問あるため、正答率が15、16番目の2項目を選び出し、その他の形式は奇数項目からなるため各1項目、合計6問題

NII-Electronic Library Service

The Japan Association of College English Teachers (JACET)

形式7項目を選び出した（表1参照）。第4～6問題形式はそれぞれ1長文について5項目あり、この場合、単に正答率の中央値をもって代表項目とすることには問題があると考えたが、幸いこれらの項目は各長文の第1項目になった。

表1　各項目の正答率

	問題形式1	問題形式2	問題形式3	問題形式4	問題形式5	問題形式6
average	0.5396	0.6161	0.7490	0.6094	0.6846	0.5691
sd	0.2099	0.2926	0.2091	0.0751	0.1774	0.1615
max	0.8255	0.8926	0.8993	0.6913	0.8591	0.7785
min	0.1208	0.1946	0.3893	0.4899	0.4430	0.3624
med	0.5436 0.5369	0.7383	0.8121	0.6242	0.7517	0.5772

次に、英語能力をこれらの問題形式（実際にはそれらの代表項目）で測定する場合の、各問題形式への最も妥当な配点比率・比重を尋ねる質問紙を作成した。実用英語検定試験をはじめ、近年の大規模テストは、採点の実施可能性と信頼性を重視する結果、客観的択一問題が多用される傾向がある。このような大規模テストの採点は通常コンピュータ処理されるため、配点比重についても、ある問題形式を1として他の問題形式を整数倍した比率で表わす必要はなく、それぞれの問題形式の比重を任意の実数（つまり少数）で表わすことが可能である。また質問紙には参考として各問題形式の項目数も記載した。

次にこの質問紙への回答を、現職の英語教員に依頼した。回収率は30％弱、回収数は98年12月末時点で合計71である。内訳は日本人54名（女性29名、男性25名）、ネイティヴ17名（女性6名、男性11名）、勤務先の学習者レベル別では日本人53名（中学校又は高等学校17名、大学27名、予備校7名）、ネイティヴ16名（中学校又は高等学校7名、大学9名）である（質問紙への未回答部分があるため欠測値ある）。

本研究発表では、回収したデータの分析をもとに以下の内容を中心に論じる。

1. 配点比重とテストの構成概念妥当性
2. 実用英語検定試験の配点比重と、英語教員が考える配点比重との異なり
3. 英語教員内の要因による配点比重の異なり
4. 英語テストの目的と配点比重のあり方
5. 英語教員の考える配点比重に基づいて行った点数計算と、実用英語検定試験の合否入れ替え率

本研究の実施には上記3名の他、加賀田哲也（大阪商業大学）、多田昌夫（大阪国際女子大学）、小山由紀江（長岡技術大学）が加わった。

NII-Electronic Library Service