

優秀修士論文概要

## 「語学力」を測るための方法および基準への探求

楊 甯

本論文は、言語テストにおける「語学力」の定義の明確化を試みるものであり、また、言語テストの結果を基にしたデータ分析を用いて、中国語教育への適用可能性を探求することを目的としている。

20世紀40年代、いわゆる「科学的測定時代」に突入する前、言語テストは言語教育および言語習得の測定道具として位置づけられていた。このような言語テストは、教育と学習の評価ツールとしての役割を果たしており、その構成内容と評価方法は、その時代の言語教育カリキュラムと教授法の影響を色濃く受けていた。このようなテストは、一般に「到達度評価 (achievement assessment)」と称されている。それに対して、特定の国や地域に依存せず、専門機関によって実施される言語テスト、すなわち受験者の語学力を評価することを主眼としたテストは「熟達度評価 (proficiency assessment)」として認識されている。このような言語テストは、受験者の学習時間や使用教材に制限がなく、さらにテストの合否が受験者の進学や就職などに大きな影響を及ぼすため、ハイ・ステークステストとも称される。本論文は、受験者の語学力を測定する言語テストに焦点を当てている。

現在、リスニング・スピーキング・リーディング・ライティングの4技能を中心に、主観テストと客観テストを組み合わせた形式で実施されている言語テストは、一般的に認知されている言語テスト形式と言えよう。これら二つの評価方法の根本的な差異は、評価の手法によるものである。客観テストは、評価結果が採点者に依存しない特性があるため、公平性や人件費の観点から多くの場面で採用されている。しかし、客観テストを通じて得られる結果が受験者の真の語学力を的確に反映しているかについては疑問視されている。特に、選択式客観テストが主観テストよりも答えやすいという点で批判されることが多い。

一方、主観テストは評価結果が採点者の主観に影響されるため、統計学が重視される現代では、その評価の変動性から計量としての取り扱いが難しいという側面がある。それにもかかわらず、文語運用や口頭表現を評価する際には、主観テストが最も効果的とされている。主観テスト評価の主観性や客観テストの制約などそれぞれの問題点に対して、受験者の語学力を的確に測定する言語テストをどのように設計するかは、言語テスト研究の重要な課題となっているであろう。

本研究では、まず、言語テストを通じて測定される語学力の定義、言語テストの現状やその進展、さらには言語テストの目的や意義に関する基本的な理論について概観する。次に、客観テストと主観テストのそれぞれの制約点を詳述し、機械学習が言語テストにどのように取り入れられるかの可能性を探る。その上で、言語テストに関する知見を基に、中国語教育に対して向上策を提言する。

本研究は序章と結論を除いて、6章からなっており、各章の内容の概要は以下である。

第1章では、一般言語学と第二言語習得の視点から見た「語学力」の定義をまとめ、その上、言語テストで測定すべき「語学力」を定義する。一般言語学における言語に関わる能力の論議の起源として、

Chomsky (1965) の考え方を無視することはできない。彼は、「理想的な話者」が所有する能力を言語能力と言語運用能力の2つに分類すべきだと主張する。この考え方はその後、Hymes (1972)、Canale & Swain (1980)、Canale (1983)、Bachman (1990) らによってさらに発展を遂げた。彼らの研究により、「言語運用能力」は「談話的能力」、「社会言語的能力」、そして「方略的能力」の3つの要素に結びつけられて認識されるようになった。「言語運用能力」に関する理論は進化を遂げたものの、それはChomsky が構築した枠組みの中での発展であると考えられる。一方、第二言語習得分野において、Faerch & Kasper (1983) は、第二言語に関する知識を「宣言的知識」と「手続き的知識」の2つ区分できるとされている。この考え方を受けている吉田 (2006) は言語の運用を個別に考えられるべきであるという立場を取っている。しかしながら、木村 (2020) は、Faerch らの主張を基に、知識の獲得過程は手続き的知識の範疇に属するという見解を示している。本論文は、これらの視点を参考にしつつ、言語テストにおいては文法知識やそれを運用する能力だけでなく、目標言語に対する認知能力の評価も含めるべきであるという立場を取る。

第2章では、言語テストに関する理解を深めるため検討する。この章は主に3つの視点で言語テストに迫る。初めに、試験実施者としての視点を採り上げる。試験実施者としては、言語テストは学習の各段階での成果を確認する手段であり、同時に受験者の総合的な語学力を評価するためのものとなっている。次に、受験者としての視点を考慮する。彼らにとって、大規模な言語テストは就職や進学といった人生の重要な節目を迎えるための手段となる。第2節では、具体的に英語と中国語の大規模な言語テストを例に、その内容構成を詳細に解説する。そして、第3節では、現代の言語テスト研究の主要な動きや大規模言語テストの最新のトレンドについて総括する。

第3章と第4章での議論の焦点は、客観テストと主観テスト（特にスピーキングテスト）がどの程度受験者の「語学力」を包括的に評価することができるか、そしてその際に直面する課題と限界についてである。これらのテストの間の基本的な違いは、評価の基準に端を発する。一言でまとめると、客観テストは採点者が誰であるかに関わらず一貫した結果を生むのに対し、主観テストの結果は採点者によって変動する可能性がある。

今日の言語評価の風景において、客観テストは受験者の語学力を評価する主要な手段として広く採用されている。しかし、その普及にもかかわらず、完全に瑕疵のない評価手段であるわけではない。客観テストが直面する最も重要な問題は、異なる難易度の問題にどのような配点を設定すれば、受験者の実際の語学力を最も公平に反映させることができるか、という点に集約される。

一方で、スピーキングテストなどの主観テストの議論では、異なる問題点が浮上する。試験官の存在が必要かどうか、そして試験官の主観に起因する評価の変動をどのように最小限に抑えるか、という問題が中心的な議論となっている。

第5章では、早稲田大学文化構想学部・文学部1年中国語春期統一試験における2020年度と2022年度の成績動向を詳細に分析する。2020年度は、新型コロナウイルスの感染拡大防止のため、学校での会場試験実施が難しくなった。その結果として、非会場型のオンラインテストへの移行が必要となり、その過程で一部の問題における未回答率が顕著に高まるという状況が生じた。それに対して、2022年度春期には「理解度確認チェック」という統一試験とほぼ同じ形式の「小テスト」が導入された。

この背景をもとに、本章では以下の3つの主要な課題に焦点を当てる。第一に、「理解度確認チェック」が統一試験全体においてどのような位置付けを持つのかについての考察である。第二の課題として、授

## 「語学力」を測るための方法および基準への探求

業形態と受験形態の双方での変化の影響を受けて、2022年度の受験者の回答傾向が、2019年度の受験者と比較してどう変動したのかを明らかにしたい。そして最後に、先に議論した客観テストの問題点への対応策として、「0-n」採点法がどれほど効果的かに関する検討を行う。

第6章では、機械学習を用いて「理解度確認チェック」および統一試験の成績を基盤に学習者のカテゴリー化を試み、それが中国語教育の現場での適用可能性について検証する。機械学習を活用した成績分析は、集団の変化傾向を中心に据える従来のアプローチとは異なり、クラスタリングを通じて、各学習者の学習意欲や特性を予測する点において独自性を持つ。

結論の部分では、本論文を通じて議論されたポイントを総括するとともに、本論文の不足や将来の課題に関して深く探求する。本論文の不足は、主として以下の3点が挙げられる。まず、「語学力」の中に内包される認知能力の具体的な定義が不足していることである。次に、主観テストの制約を検討する上で、ライティングテストが十分に考慮されていないことである。最後に、機械学習を用いた分析において、「理解度確認チェック」と統一試験の関連性のみを焦点を当て、他の変数を考慮しきれていないことである。これらの不足を踏まえ、今後は「語学力」を的確に評価する言語テストの理想的なモデルの探求を目指したい。